The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps

October 13, 2005

Abstract

This paper studies the determinants of children's scores on tests of cognitive achievement in math and reading. Using rich longitudinal data on test scores, home environments, and schools, we implement alternative specifications for the production function for achievement and test their assumptions. We do not find support for commonly used restrictive models that assume test scores depend only on contemporaneous inputs or that assume conditioning on a lagged score captures all the effects of past inputs. Instead, the results show that both contemporaneous and lagged inputs matter in the production of current achievement and that it is important to allow for unobserved child-specific endowment effects and endogeneity of inputs. Using a specification that incorporates these features, we study the sources of test score gaps between black, white and Hispanic children. The estimated model captures key patterns in the data, such as the widening of minority-white test score gaps with age and differences in the gap pattern between Hispanics and blacks and between boys and girls. We find that equalizing home inputs at the average levels of white children would close the black-white test score gap by about 25% and close the Hispanic-white gap by about 30%.

JEL Codes: J24, J15, I20

Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104

1 Introduction

It is well documented that scores on cognitive tests taken during adolescent years are correlated with adult labor market outcomes, such as educational attainment and earnings.[1] Even scores on tests taken as early as age seven have been shown to be correlated with measures of labor market success.[2] These findings have led many researchers to assign a large role to "premarket factors" in explaining earnings inequality, where premarket factors are broadly interpreted to represent endowed ability, the effects of family background and the influence of schools.

Premarket factors are also considered an important part of the explanation for racial differences in test score performance and labor market outcomes.[3] While it is conceivable that test score gaps could arise from differential investment in children based on expectations about future labor market returns (a post market rather than a premarket factor), Carniero, Heckman and Masterov (2002) argue that this is an unlikely explanation for gaps observed for children at the ages of school entry. Test score gaps between white and black children emerge at early ages and tend to widen with age.[4] Overall, the average test score gap between whites and minorities has narrowed substantially since the 1970's, but black children still score about 15-25% lower than whites on average and Hispanic children about 10% lower.[5]

The belief that eliminating racial differences in test score performance would reduce

[1] See e.g. Leibowitz (1974), Murnane, Willett and Levy (1995), Neal and Johnson (1996), Keane and Wolpin (1997), and Cameron and Heckman (1998).

[2] For example, Robertson and Symons (1996) find that age seven test scores predict occupational choices; Currie and Thomas (1999) document their correlation with adult educational and labor market outcomes. Hutchinson, Prosser and Wedge (1979) observe that test scores at age seven are highly correlated with scores at age 16. All of these studies are based on data gathered by the British National Child Development Survey, which has been following a cohort of Children born during one week in March, 1958.

[3] Neal and Johnson (1996).

[4] See Carniero, Heckman and Masterov (2002), Levitt and Fryer (2002), Phillips, Crouse and Ralph (1998), and section 3 of this paper for evidence on the widening of cognitive achievement scores by age. There is, however, some debate over whether test scores widen as children progress through school grades (Ludwig (2003)). Carniero and Heckman (2003) discuss gap patterns by age in noncognitive test score measures.

[5] See Jencks and Phillips (1998), Cook and Evans (2000) for a discussion of trends in scores on NAEP (National Assessment of Educational Progress) tests. Hedges and Nowell (1998, 1999) analyze data from six surveys that include EEO (Equal Educational Opportunity Data), National Longitudinal Study of the High School Class of 1972, High School and Beyond, National Longitudinal Survey of Youth (1979), National Education Longitudinal Study of 1988, and NAEP.

inequality in labor market outcomes is an important motivation for an extensive, multidisciplinary literature aimed at understanding the determinants of children's test scores. One branch of the literature studies the role of parental characteristics and the early home environment in producing early cognitive skills. Another branch examines the influence of school characteristics on children's test scores.

In both literatures, there have been debates over which inputs increase children's achievement and to what extent. For example, many child development studies consider the question of whether early maternal employment is detrimental for children's achievement. There is wide variation in reported empirical estimates, even for studies based on the same data.[6] In studies of school effects, there are disagreements over whether inputs such as pupil-teacher ratios, teacher experience and teacher salaries matter in producing cognitive skills.[7] A leading candidate for explaining why studies reach such different conclusions is that the statistical models used to estimate relationships between inputs and outcomes are misspecified. (Krueger, 2003, and Todd and Wolpin, 2003)

Ideally, in analyzing cognitive achievement of children, it would be useful to have access to data on all past and present family and school inputs as well as information on children's heritable endowments. No dataset is that comprehensive, so researchers have had to confront problems of missing data and imprecisely measured inputs. Datasets used in studies of early childhood development often have information on family inputs but lack information on schools.[8] Datasets used in studies of school effects often contain information on contemporaneous school inputs, but have limited or no information on the home environment and on historical school inputs.

Confronted with what are sometimes severe data limitations, researchers have pursued a variety of estimation strategies to overcome them. One approach explicitly recognizes the presence of omitted variables and develops estimators that allow for them. For example,

[6] For example, estimates range from maternal employment being detrimental (Baydar and Brooks-Gunn, 1991; Desai et. al., 1989; Belsky and Eggebeen, 1991), to its having no effect (Blau and Grossberg, 1992) to its being beneficial (Vandell and Ramanan, 1992).

[7] For example, Krueger (1998) and Hanushek (1998) both analyze data on National Assessment of Educational Progress (NAEP) test scores, with Krueger concluding that increases in per pupil expenditure have led to modest gains and Hanushek concluding no effect. Also, see a summary of the issues in the debate surrounding the impact of school quality on achievement and earnings in Burtless (1996).

[8] For example, Baharudin and Luster (1998) and Crane (1996) analyze effects of family inputs on cognitive achievement of school-age children without taking into account the contribution of schools.

Murnane, Maynard and Ohls (1981) use school fixed effects to address the problem of missing school inputs, under the assumption that children within the same school receive the same inputs. Rosenzweig and Wolpin (1994) and Altonji and Dunn (1996) address the same problem using sibling fixed effects. An alternative approach that is commonly taken when the data lack information on historical input measures, is to adopt a value added specification, which assumes that a previous test score serves as a sufficient statistic for the influence of historical inputs. Another remedy to the missing data problem is to use one or more proxy variables that are not considered direct inputs into cognitive achievement, but are included in the analysis under the presumption that they alleviate omitted variables bias given their correlation with omitted inputs. Variables such as family income or race could be considered such proxy variables.[9]

This paper studies the determinants of children's scores on tests of cognitive achievement in math and reading. Specifically, we estimate a production function for achievement that is consistent with theoretical notions that child development is a cumulative process depending on the history of family and school inputs and on heritable endowments. We base our analysis on rich data containing longitudinal information on both family and school inputs. These data enable implementation of more general models and allow testing of many of the modeling assumptions commonly invoked in the literature. Our work builds in some ways on Boardman and Murnane (1979), who were the first to formalize a cumulative model of the cognitive achievement production function and to discuss its implementation in cross-section and panel data settings.[10] This paper also builds on Todd and Wolpin (2003), which surveyed various approaches in the literature to estimating the cognitive achievement production function.[11]

This paper has two main goals: to quantify the impact of home inputs and school inputs on children's achievement and to analyze the contribution of home and school inputs in accounting for racial test score gaps. Our analysis is based on data from the National Longitudinal Surveys of Labor Market Experience - Children Sample (NLSY79-CS) merged together with school data obtained from three sources: the Common Core Data (CCD), the

[9] See Todd and Wolpin (2003) and section four below for a discussion of potential biases associated with the use of proxy variables.

[10] We recently became aware of this insightful but overlooked paper.

[11] For each method, they presented the identifying assumptions, the data requirements, and conditions under which assumptions of the estimation method could be tested.

School and Staffing Survey (SASS) and the American Federation of Teachers (AFT). The NLSY79-CS data contain detailed longitudinal information on children's home environments and on child achievement as measured by scores on tests that are administered biannually. Although these data are also deficient in some respects, they come closest to the ideal for estimating cognitive achievement production functions. Using these data, we implement alternative specifications and perform a number of model specification tests aimed at assessing the importance of missing data on inputs, endogenous inputs and unobserved heterogeneity, all of which we find to be empirically relevant. Evidence based on a specification that allows for these features of the data finds both current and past home inputs to be significant determinants of test score outcomes. The effects of the school input variables on test scores are usually not precisely measured in the specifications that allow for unobserved heterogeneity.

We use our estimates of the cognitive achievement production function parameters to examine the extent to which home and school input differences can account for racial disparities in test scores among African American, white and Hispanic children. Our work differs from earlier studies, in part, because our specification allows for unobserved endowment effects, potentially endogenous input choices, and for the cumulative effects of lagged inputs.[12] The empirical results show that equalizing home input levels at the average level observed for white children would close about 25% of the black-white test score gap (in both math and reading) and 30% of the Hispanic-white test score gap. We also find that the estimated cognitive achievement production function fits well the pattern of rising black-white test score gaps with age as well as differences in test score gap patterns between girls and boys.

Our finding that home input gaps are important in accounting for racial test score gaps contrasts with findings reported in Levitt and Fryer (2004). That paper argues that home input gaps cannot account for black-white test score gaps, because home input gaps remain roughly constant over time whereas test score gaps widen with age. Their observation is based on a specification in which test scores depend on current home inputs. In this paper, we show that a specification that allows test scores to depend on both current and

[12] For example, Cook and Evans (2000) decompose test score differences into components due to changing relative levels of parental education, changing levels of school quality, and a narrowing of within school gaps, using data from NAEP (National Assessment of Educational Progress). The NAEP dataset has the advantage of being a large representative sample with multiple observations per school, but it contains little information on childrens' home environments and is not longitudinal. Also, see Fuchs and Reklis (1994) for an analysis of the sources of racial math test score differences using state-level NAEP data.

historical home inputs can explain why black-white test score gaps widen with age, because a constant home input gap over time implies a widening cumulative gap. In a recent working paper, Levitt and Fryer (2005) appear to dismiss the cumulative specfication as a potential explanation for the widening test score gap based on a finding that the estimated coefficients associated with current home inputs do not change much when lagged inputs are included in the specification.[13] However, this finding really has no bearing on whether home inputs can account for the widening of the gap, which depends not on whether the estimated current input effect is changed but rather on whether lagged inputs matter in the specification.

This paper proceeds as follows. Section two of the paper proposes a conceptual framework for modeling the cognitive achievement production function and considers its empirical implementation. Section three describes our data sources and the variables used to represent home and school inputs into the production process. Section four presents estimates of the cognitive achievement production function obtained under alternative specifications. Our specification tests reject all of the more restrictive specifications in favor of one that allows for endogenous inputs and unobserved endowments. Section five uses the estimated cognitive achievement production function to evaluate the sources of racial disparities in test scores. Section six concludes.

2 Alternative Approaches to Modeling and Estimating the Production Function for Achievement

2.1 A General Framework

In this section, we lay out a general framework for modeling the cognitive achievement production function. Let T_{ija} be a test score measure of achievement for child i residing in household j at age a. We conceive of knowledge acquisition as a production process in which current and past inputs are combined with an individual's genetic endowment of mental capacity (determined at conception) to produce a cognitive outcome.[14] We distinguish be-

[13] Levitt and Fryer (2005), p. 17.

[14] The production function framework was first formally modeled by Ben Porath (1967) in the context of an individual decision-maker choosing the level of (time and money) resources to devote to human capital investments. It has since served as the basis for much of the literature on skill acquisition in economics. Leibowitz (1974) was the first to extend this conception to home investments in children.

tween two kinds of inputs into the production function: inputs that are endogenous and reflect choices made by parents (such as how often parents read to the child or the characteristics of the school the child attends) and inputs that are exogenous and not subject to parental choice (such as a random illness or the quality of the teacher the child is assigned to within the school).

Denoting the vector of parent-chosen inputs at a given age as Y^c_{ija}, and exogenous inputs as Y^e_{ija} and the vectors of their respective input histories as of age a as $Y^c_{ij}(a)$ and $Y^e_{ij}(a)$, and also denoting a child's endowed mental capacity ("ability") as μ_{ij0}, the achievement production function is given by

$$T_{ija} = T(Y^c_{ij}(a), Y^e_{ij}(a), \mu_{ij0}, \varepsilon_{ija}), \tag{1}$$

where ε_{ija} represents measurement error.[15]

As described in the introduction, the empirical implementation of (1) is difficult for three reasons: (i) heritable endowments are unobservable; (ii) data sets on inputs are incomplete (i.e. have incomplete input histories and/or missing inputs) and; (iii) inputs may be chosen endogenously with respect to unobserved endowments and/or prior realizations of achievement.

2.2 The Contemporaneous Specification

Suppose that the dataset only contained information on contemporaneous input measures and all historical data were missing, as is typical with cross-section data. In this case, one option is to adopt a *contemporaneous specification,* which relates an achievement test score measure solely to contemporaneous measures of inputs:

$$T_{ija} = T(Y^c_{ija}, Y^e_{ija}) + \varepsilon'_{ija}, \tag{2}$$

where ε'_{ija} is an additive error and where, for now, we assume that all of the contemporaneous inputs are observed.[16] The following assumptions on the production technology and on the input decision rules would justify the application of (2) as a way of estimating (1).

[15]The $T(\cdot)$ function is assumed not to differ depending on the age of the child. See Todd and Wolpin (2003) for a discussion of the case where it may vary.

[16]The problem of missing data on inputs will be considered below.

(i) Only contemporaneous inputs matter to the production of current achievement.

or

(ii) Inputs are unchanging over time, so that current input measures capture the history of inputs.

and, in addition to (i) or (ii),

(iii) Contemporaneous inputs are unrelated to (unobserved) endowed mental capacity.

In a contemporaneous specification, the residual term ε'_{ija} includes any omitted factors—the history of past inputs, endowed mental capacity–as well as measurement error. The assumptions necessary to consistently estimate the impact of contemporaneous inputs, the only observable data, are obviously quite severe.

In particular, assumption (iii)–that inputs and endowed ability are uncorrelated–is inconsistent with economic models of optimizing behavior. Economic models in which parents care about a child's cognitive development imply that the amount of resources allocated to the child, in the form of purchased goods and parental time, will be responsive to the parent's perception of a child's ability. Thus, while the contemporaneous specification has weak data requirements, strong assumptions are required to justify its application.[17]

2.3 The Value-Added Specification

An alternative specification that is sometimes adopted when data on input histories are not available is a *value-added model*. In its most common form, this specification relates an achievement outcome measure to contemporaneous school and family input measures and a lagged (baseline) achievement measure. The baseline achievement measure is taken to be a sufficient statistic for input histories as well as for the unobserved endowed mental capacity. Evidence based on the value-added specification is generally regarded as being better (more convincing) than that based on a contemporaneous specification. (See, e.g., Summers and Wolfe, 1977, Hanushek, 1996).

Let X_{ija} denote the vector of inputs that are observed in the data and υ_{ija} the inputs that are not observed. The conventional value-added specification assumes that equation (1) can be written as an additively separable function only of a previous period baseline test score and observed contemporaneous inputs (inputs applied between the baseline measure

[17] The test score specification estimated in Fryer and Levitt (2002) can be viewed as a form of the contemporaneous specification. It relates current test scores to contemporaneous measures describing the child's home and school environment, without allowing for endogeneity of inputs or for the effects of unobserved endowments.

and a current measure):[18]

$$T_{ija} = X_{ija}\alpha + \gamma T_{ij,a-1} + \eta_{ija}. \tag{3}$$

As first shown in Boardman and Murnane (1979), a value-added specification imposes severe restrictions on the coefficients of the cumulative model.[19] In our context, consider the regression analog of (1), where test scores can depend on both observed and unobserved contemporaneous and historical inputs and on unobserved endowments:

$$\begin{aligned} T_{ija} &= X_{ija}\alpha_1 + X_{ija-1}\alpha_2 + ... + X_{ij1}\alpha_a + \\ &\quad \beta_a \mu_{ij0} + \{\upsilon_{ija}\rho_1 + \upsilon_{ija-1}\rho_2 + ... + \upsilon_{ij1}\rho_a + \varepsilon_{ija}\} \\ &= X_{ija}\alpha_1 + X_{ija-1}\alpha_2 + ... + X_{ij1}\alpha_a + \beta_a \mu_{ij0} + \epsilon_{ija}. \end{aligned} \tag{4}$$

Here, ϵ_{ija} is an error term that includes the effect of the history of unobserved inputs and measurement error, so ϵ_{ija} would be expected to be serially correlated. Subtracting $\gamma T_{ij,a-1}$ from both sides of (4) and collecting terms gives,

$$\begin{aligned} T_{ija} &= X_{ija}\alpha_1 + \gamma T_{ij,a-1} + X_{ija-1}(\alpha_2 - \gamma\alpha_1) + ... + X_{ij1}(\alpha_a - \gamma\alpha_{a-1}) \\ &\quad +(\beta_a - \gamma\beta_{a-1})\mu_{ij0} + \{\epsilon_{ija} - \gamma\epsilon_{ij,a-1}\} \end{aligned} \tag{5}$$

where $\epsilon_{ija} - \gamma\epsilon_{ij,a-1} = \upsilon_{ija}\rho_1 + \upsilon_{ija-1}(\rho_2 - \gamma\rho_1) + ... + \upsilon_{ij1}(\rho_a - \gamma\rho_{a-1}) + \varepsilon_{ija} - \gamma\varepsilon_{ija-1}$. For (5) to reduce to (3), three conditions suffice:

(i) Coefficients associated with observed inputs geometrically (presumably) decline with distance, as measured by age, from the achievement measurement and the rate of decline is the same for each input, (i.e. $\alpha_j = \gamma\alpha_{j-1}$ for all j).

(ii) Condition (i) also holds for omitted inputs ($\rho_j = \gamma\rho_{j-1}$ for all j) and the contemporaneous omitted input υ_{ija} is uncorrelated with included inputs and with the baseline test score; or omitted inputs (current and lagged) are uncorrelated with included inputs and with the baseline test score.

(iii) The impact of the endowment geometrically declines at the same rate as input effects, i.e., $\beta_a = \gamma\beta_{a-1}$.

For the ols estimator of α_1 to be consistent, ϵ_{ija} must also be serially correlated and the degree of correlation must exactly match the rate of decay of input effects (so that $\eta_{ija} =$

[18] A more restrictive specification sometimes adopted in the literature sets the parameter on the lagged achievement test score to one ($\gamma = 1$) and rewrites (3) as

$$T_{ija} - T_{ij,a-1} = X_{ija}\alpha + \eta_{ija},$$

which expresses the test score gain solely as a function of contemporaneous inputs.

[19] See also Todd and Wolpin (2003).

$\epsilon_{ija} - \gamma\epsilon_{ija-1} = \upsilon_{ija}\rho_1 + \varepsilon_{ija} - \gamma\varepsilon_{ija-1}$ is an iid shock). If this condition is not satisfied, then baseline achievement, T_{ija-1},will be correlated with η_{ija}. (Baseline achievement is necessarily correlated with its own measurement error (ε_{ija-1}) and may also be correlated with omitted inputs υ_{ija}.)

If we drop the assumption that the impact of the mental capacity endowment declines at the same rate as the decay in input effects (given above by (iii)), then the error in (3) would include the endowment, *i.e.*, assuming that $\beta_a - \gamma\beta_{a-1} = \beta'$ is a constant independent of age, yields

$$T_{ija} = X_{ija}\alpha + \gamma T_{ij,a-1} + \beta'\mu_{ij0} + \eta_{ija}, \tag{6}$$

instead of (3).

Estimation of (6) by ols is problematic. As with the contemporaneous specification, one requirement for ols to be consistent is that contemporaneous inputs and unobserved mental capacity be orthogonal. However, even if that orthogonality condition were satisfied, ols estimation of (6) would still be biased, because baseline achievement must be correlated with endowed mental capacity. If the endogeneity is not taken into account, then the resulting bias affects not only the estimate of γ, but may be transmitted to the estimates of all the contemporaneous input effects. Thus, the value-added specification does not easily accommodate the presence of unobserved endowment effects.

2.4 The Cumulative Specification

When data are available on historical inputs, it is possible to consider direct estimation of the cumulative specification (4). In the discussion that follows, we assume that any omitted inputs and measurement error in test scores are uncorrelated with included inputs. Under this assumption, the challenge in estimating (4) is that behavior in the choice of inputs may induce correlations between the observable inputs and unobserved child endowments. A class of estimators used to “control” for permanent unobservable factors in estimating (4) makes use of variation across observations within which the unobservable factor is assumed to be fixed. Two such “fixed effect” estimators prominent in the literature use variation that occurs within families (across siblings) or within children (at different ages).

Within-family estimators exploit the fact that children of the same parents have a common heritable component. In particular, assume that endowed mental capacity can be decomposed into a family-specific component and may have a child-specific component, denoted as μ_0^f and μ_0^c. Thus, siblings have in common the family component, but have their own individual-specific child components. Rewriting (4) to accommodate this modification yields

$$T_{ija} = X_{ija}\alpha_1 + X_{ija-1}\alpha_2 + ... + X_{ij1}\alpha_a + \beta_a\mu^f_{ij0} + \beta_a\mu^c_{ij0} + \epsilon_{ija}. \tag{5'}$$

Now, suppose that longitudinal household data on achievement test scores and on current and past inputs are available on multiple siblings, as in the NLSY79-CS data that we use in the empirical work We distinguish between two types of data. In the first case, data are available on siblings at the same age.[20] In the second case, data are available on siblings in the same calendar year, which generally means that they will differ in age.

Consider the estimator in the case of two siblings, denoted by i and i' observed at the same age a. Differencing (4) yields

$$T_{ija} - T_{i'ja} = (X_{ija} - X_{i'ja})\alpha_1 + ... + (X_{ij1} - X_{i'j1})\alpha_a + [\beta_a(\mu^c_{ij0} - \mu^c_{i'j0}) + \epsilon_{ija} - \epsilon_{i'ja}] \tag{7}$$

In estimation the residual term will include all the terms within the square brackets. Consistent estimation of input effects by ols, therefore, requires that inputs associated with any child not respond either to the own or sibling child-specific endowment component.

Furthermore, given that achievement is measured for each sibling at the same age, the older child's achievement observation (say child i) will have occurred at a calendar time prior to the younger sibling's observation. Thus, the older sibling's achievement outcome was known at the time input decisions for the younger child were made, at the ages of the younger child between the older and younger child's achievement observations. Thus, consistent estimation of (7) by ols also requires that input choices are unresponsive to prior sibling outcomes (otherwise the realizations of ϵ_{ija} will affect some of the inputs to sibling i'). In essence, this estimation procedure is justified when intra-household allocation decisions are made ignoring child-specific endowments and prior outcomes of all the children in the household (Rosenzweig, 1986).

Within-child estimators are feasible when there are multiple observations on achievement outcomes and on inputs for a given child at different ages.[21] Consider differencing the achievement technology at two ages, a and $a-1$,

$$\begin{aligned} T_{ija} - T_{ija-1} &= (X_{ija} - X_{ija-1})\alpha_1 + (X_{ija-1} - X_{ija-2})\alpha_2 + ... + \\ &\quad (X_{ij2} - X_{ij1})\alpha_{a-1} + X_{ij1}\alpha_a \\ &\quad + [\beta_a - \beta_{a'}]\mu_{ij0} + \epsilon_{ija} - \epsilon_{ija-1}. \end{aligned} \tag{8}$$

[20] Notice that unless the siblings are twins, the calendar time at which achievement measures are obtained must differ.

[21] The within-family estimator based on siblings of different ages can be viewed as a special case of the within-child estimator based on test scores of the same child measured at different ages.

The parameters of (8) can be consistently estimated under the following assumptions. The first is that the impact of the capacity endowment on achievement must be independent of age ($\beta_a = \beta_{a'}$), in which case differencing eliminates the endowment from (8). In that case, orthogonality between input choices and capacity endowments need not be assumed. However, because any prior achievement outcome is known when later input decisions are made, it is necessary to assume that later input choices are invariant to prior own achievement outcomes.

Consider relaxing the assumption that input choices do not respond to prior realizations of achievement. If the shocks in (8) result from unforeseen exogenous factors (*e.g.*, a random illness or randomly drawing a bad teacher) and if the impact of these factors on achievement has limited persistence, then input levels prior to the earlier achievement observation can serve as instrumental variables for the differenced inputs in estimating (8). For example, if the achievement tests used in the differenced estimation procedure are taken at ages 8 and 7, then the set of inputs at ages earlier than 3 could serve as instruments. However, there are more parameters in (8) than instruments–at least as many as the number of measured inputs–so identification cannot be achieved with these orthogonality conditions alone. We can augment the set of instruments to include inputs associated with the child's siblings applied at a time sufficiently prior to the earliest observation used to implement the within-child estimator.[22] This is the strategy taken in section 4 where we implement both within-sibling and within-child estimators, with and without controlling for endogeneity of input choices.

Finally, none of the IV approaches are valid if omitted inputs are not orthogonal to the included ones, because omitted inputs that are correlated with observed input choices (presumably because they also reflect choices) are also likely to be correlated with the instrumental variables. It is therefore important that the data contain a large set of inputs spanning both family and school domains to make plausible the required assumption that omitted inputs are exogenous.

3 Data

As described in section two, the data requirements for implementing the cumulative specification of the cognitive achievement production function are demanding. A researcher needs a complete history of inputs, beginning at the child's conception, including both those that are

[22] This kind of informational constraint was used by Rosenzweig and Wolpin (1988, 1995) to estimate birth weight production functions.

subject to parental choices and that are exogenous. In addition, to account for unobserved endowments one needs multiple observations on achievement measures, either for siblings at the same ages or for the same child at different ages. Although there does not exist a data set that satisfies all these requirements, the NLSY79 Child Sample (NLSY79-CS) comes closest to the ideal.

The NLSY79-CS is a sample of all children ever born to the women respondents of the NLSY79. The NLSY79 is itself a nationally representative sample of individuals who were age 14-21 as of January 1, 1979, with significant oversamples of blacks and Hispanics. The survey collects extensive information about schooling, employment, marriage, fertility, income, assets, alcohol and drug use, participation in public programs and other related topics, many as event histories. For example, employment events are known up to the week, marriage and fertility events to the day and school enrollment to the month. This enables the researcher to create an almost complete life history for each respondent for many important events dating back to age 14.

Beginning with the 1986 interview, a separate set of questionnaires were developed to collect information about the cognitive, social and behavioral development of the children of the NLSY79 respondents. Questionnaires were administered to the women (cum mothers) of the children as well as to the children themselves. These interviews have been conducted biannually since 1986. By 2000, the most recent survey data publicly available, over 11,000 children were interviewed. Approximately 28 percent of the children in 2000 were African American, 19 percent Hispanic and the rest mostly white.

Cognitive Achievement Measures Our analysis restricts attention to two cognitive tests that were administered to all children starting at age five: the Peabody Individual Achievement Test in mathematics (PIAT-M) and the Peabody Individual Achievement Test in reading recognition (PIAT-R). The PIAT tests are designed to measure academic achievement. They were administered each year of the survey, and many of the children in the sample have two or more scores. Completion rates for the PIAT's have been around 90 percent.

Table 1 shows the average PIAT Math and Reading scores by race/ethnicity. In our work, we use raw test scores rather than age-adjusted or normed percentile scores, because we want an absolute measure of achievement that captures gains over time as additional input investments are made in a child. The average raw scores for African American and Hispanic children are about 5 points lower than the average score for white children, a gap of 12%.

Figures 1a and 1b plot the average PIAT Reading and Math test scores by age, by gender and by race/ethnicity. The lower panel shows the black-white and Hispanic-white gap. At age six, there is only a one point gap in reading scores for black children, which is smaller than the two point gap observed for Hispanic children. However, the gap widens over the first two years of school and, by age eight, the black-white reading gap is 6 in comparison with a gap of 4 for Hispanic children. The gap continues to grow through age 12, although more slowly. As seen in the plots for boys and girls, the widening of the black-white test score gap is most pronounced for boys. The white-Hispanic gap also widens, but to a lesser extent, and reaches a peak of 5.5 at age 12. For both the black and Hispanic children, there is some evidence of convergence in the reading score gap between ages 12 and 13.

The patterns for PIAT-math scores are similar, also exhibiting a widening gap that is especially pronounced for black boys. A major difference between the reading and math score patterns is that the gap emerges at an earlier age for math, but stabilizes after age eight. At age six, minority children already have on average a 4 point lower math score. As seen in the figure, the black-white gap tends to be smaller for black girls than for black boys but similar by gender for Hispanics.

Home Input Measures The NLSY79-CS includes a battery of questions about the home environment of the child called the Home Observation Measurement of the Environment-Short Form (HOME-SF).[23] The HOME-SF consists of four different instruments that depend on the age of the child: ages 0-2, 3-5, 6-9 and 10 and above. The instrument is (self-administered) to the mother of the child. A second version is filled out by the interviewer. Researchers can use either the answers to individual items or scales provided in the public use files. The total raw score is a simple summation of responses (modified so each has a {0,1} domain}) of individual items.

Some of the items in the home can be directly linked to cognitive achievement in the sense that they are related to learning-specific skills. For example, mothers of children under the age of 10 are asked how often they read stories to their child, and mothers of children between the ages of 3 and 5 are asked whether they help their child to learn numbers, the alphabet, colors or shapes and sizes. Other items are not so easily tied to cognitive achievement, but may be thought of as creating an environment conducive to learning. For example, mothers are asked how many books the child has, whether the family encourages the child to start

[23] As the name suggests, the short form is a modification of a version that is about twice as long. The HOME was created by Caldwell and Bradley (1984). Some parts of the shortened version used in the NLSY79-CS were created by them and all were reviewed by them. The HOME (-SF) is widely used and there exists considerable research on the validity and reliability (see the citations in the 1996 *Users Guide*).

and keep doing hobbies, and whether the family takes the child to museums and/or theatrical performances.

In the empirical work reported below, we use the home scale provided in the public use files as our measure of the home input. As described in section two, we consider both current home inputs and historical home inputs as potential determinants of current test scores. The variables we use are shown in Table 1. *Current Home Score* gives the contemporaneous home input score that is measured at the same time the PIAT test is administered. The variable *Lag Home Score* is the home input measure obtained at the previous survey round, for children age 6-13. The variable *Lag Lag Avg* is the average of any still earlier home score input measures, obtained in earlier survey rounds for children in the same age interval.[24] The variables *Home Age 0-2* and *Home Age 3-5* give the home scores at earlier ages.[25] As seen in Table 1, the home score input measures for blacks and Hispanics are about 15% lower than for whites over all age ranges.

Figure 2 plots the current home score by age, gender and race/ethnicity. The plots show that the gap in home scores (relative to whites) is similar for blacks and Hispanics, and remains roughly constant across ages. The plots by gender show that black boys have slightly lower home scores than Hispanic boys, but the reverse is true for girls.

To get a better idea of what the home scale measures, Tables A.1-A.4 in appendix A compare the average scores by race/ethnicity for the individual items of the cognitive home scale for children in different age ranges.[26] About 2/3 of the items in the home scale are based on mother self-reports of her own and her child's activities and about 1/3 of the items correspond to interviewer observations about the child's home environment. The average scores for the African American and Hispanic mothers tend to be similar and tend to be lower than the scores for white mothers for most of the individual items. The differences are particularly notable for the questions related to number of books in the child's possession, the number of times the mother reads to the child, and the teaching activities the mother engages in with the child. For example, 94% of white mothers report that their age 3-5 toddler has 10 or more books in comparison with 57% of black mothers and 63% of Hispanic mothers. The difference in book ownership persists for children in all the age ranges. 70%

[24] Given the biannual nature of the survey and the change in the format of the home scale before and after age 6 (see Appendix Tables A.1-A.4), we do not consider the home scales prior to age 6 as representing lags of home scales measured at age 6 or later. Therefore, the Lag Home Score exists only for children age 8 and older and the double lag average home score for children age 10 and older.

[25] If multiple home scores are available within each of the age intervals (0-2 or 3-5), then we use an average measure.

[26] The questions that are asked of the mother differ slightly across four different age ranges.

of these same mothers report reading stories to their toddler at least 3 times a week, in comparison with 40% of black mothers and 44% of Hispanic mothers. 66% of black mothers and 70% of Hispanic mothers report teaching their age 3-5 child numbers in comparison with 78% of white mothers. For older children age 6-9, 61% of white children receive special lessons or participate in organizations that encourage sports, arts, dance or drama, compared to 41% for black children and 39% for Hispanic children. The items of the home scale based on interviewer observations also show some differences by race/ethnicity, but they tend to be smaller than the differences observed on the self-report items. Thus, examination of the individual items of the home scores reveals some stark racial/ethnic differences for children in all the age ranges, especially for the items that are self-reported by the mother related to books, reading and teaching activities.

Maternal Characteristics Because of the sample design of the NLSY79-CS, there is essentially continuous time data on maternal employment. Summary measures are provided in the public use data that have matched the event history employment data of the NLSY79 women respondents to each child's lifetime. These measures include weeks and hours worked in each of the first twenty quarters after the birth of the child. We interpret mother's employment is an input measure, because it represents time that is unavailable to spend with the child.[27] As seen in Table 1, labor force participation for mothers with young children is much higher among white than among black and Hispanic mothers. About half of African American and Hispanic mothers worked when their children were less than one year of age, in comparison with about two-thirds of white women.

In addition to information on employment, information is also available on mother's completed schooling, which is updated in each year in which the mother attended school. Because some women return to school after having children, both within-family and within-child estimators can be used to estimate maternal schooling effects on children's achievement.[28] Women with higher school attainment presumably have more knowledge to transmit to their children and/or may be better teachers. A comparison of mothers' schooling levels by race/ethnicity shows that white mothers have the highest average years of schooling (12.9), African American mothers the second highest (12.3), and Hispanics the lowest (11.3).

In addition to schooling, the NLSY79 also contains a measure of ability for the mothers, their score on the Armed Forces Qualifying Test (AFQT). A direct measure of mother's

[27] It would be preferable to have a direct measure of time the mother spends interacting with the child, but we take time spent not working as a proxy.

[28] Rosenzweig and Wolpin (1994) exploit the interruption in schooling that occurred for some NLSY79 mothers.

knowledge is a potentially important factor in the production of cognitive skills in children. As seen in Table 1, the AFQT score for white mothers is close to the median, while the average percentile rank for African American and Hispanic mothers, 21 and 25, is much lower.[29]

The NLSY79 includes only limited information about fathers. In fact, identifying the biological father is problematic. Although the public-use data include a variable indicating presence of the biological father in the household, the variable is missing in many cases.

Child Characteristics In addition to standard information on race and gender of the child, the NLSY79-CS also contains information on other characteristics that are potential determinants of a child's cognitive achievement, such as birth order and birthweight.[30] As shown in Table 1, African American children have on average lower birthweight than white or Hispanic children. The disparity of about 6-8 ounces is due either to biological factors or to differences in prenatal investments. White children are more likely to be first or second born because white women have fewer children.

School Inputs The major weakness of the NLSY79-CS is the paucity of data on schools. Implementing the cumulative model described by equation (1) in the previous section requires both contemporaneous and historical data on school inputs. We therefore obtain schooling data from other sources that we merge with the NLSY79-CS data using information on the child's grade, county and state of residence, and whether the child was attending private school.[31]

One of the data sources we use is the Common Core Data (CCD), which contains information on all public schools and on the characteristics of students both at the school and district levels. In the CCD, schools report the number of full-time equivalent teachers and the number of pupils enrolled, which we use to calculate pupil-teacher ratios for each school. Because elementary grades and upper level grades are usually offered in separate schools, we obtain separate pupil-teacher ratio averages for grades 1-6 and grades 7-12. We constructed both county level and state level pupil-teacher ratio variables, which we merged with the NLSY79-CS data.

A limitation of the Common Core Data is that it contains relatively little information

[29]The racial disparities in mother's AFQT scores measured using the child as the level of observation is much greater than the disparity measured using the mother as the unit of observation.

[30]See, e.g., Rosenzweig (1986).

[31]County and state of residence are available at each survey round of the NLSY79 respondents (and their children) and can be obtained as a restricted data file from the Bureau of Labor Statistics.

on characteristics that may be related to the quality of teachers, such as their years of teaching experience or their salaries. We therefore combined information from the CCD with additional information from the School and Staffing Survey (SASS) and from the American Federation of Teachers. The SASS which provides richer information on schools than does the CCD, but for a smaller sample of schools (14,500 schools are sampled in 1999, which is approximately 1 in 7 schools) and for a subsample of years covered by the NLSY79-CS data.[32] To avoid having too many missing observations, we aggregated the SASS data to the state level, and then, as with the CCD, match it to the children in the NLSY79-CS. From the American Federation of Teachers, we obtained a series of average teacher salaries by state, for the years 1984-2001.

The schooling inputs on which we focus in the analysis are pupil-teacher ratios and teacher salaries. We also estimated specifications using data on teacher's education, teacher experience, hours/week spent teaching math and English (separately), and teacher certification. These variables do not appear in the final specifications as inputs, because estimates of their effects were never precise. We therefore adopted a more parsimonious specification that includes two conventional measures of school inputs: pupil-teacher ratios and teacher salaries.

Table 1 shows average pupil teacher ratios and average teachers' salaries for white, African American and Hispanic children, where the average is taken over the child's school history for the years in which the school input measures are available.[33] Although historically, African American children attended schools that were of much lower quality than white children, there has been substantial convergence in empirical measures of schooling quality over time. Boozer, Krueger, and Wolken (1992) note that in 1970 the pupil-teacher ratios in schools attended by black children were on average 11% higher than in schools attended by whites, but by 1990 there was no difference.[34] In our schooling data, the average pupil-teacher ratios are lowest for African American children and highest for Hispanic children. Teacher salaries on average are highest for Hispanic children and lowest for African American children.

[32] SASS data is available for a subset of years in 1987-1994 and 1999-2000.

[33] We attempted to construct separate contemporaneous and lagged average measures of school inputs (as with the home inputs), but, perhaps due to the higher level of aggregation, there was substantial colinearity and we were unable to obtain precise estimates of their separate effects. Therefore, we use one cumulative measure.

[34] See also Card and Krueger (1992) for evidence on the convergence of schooling quality in black and white schools over the last century and an analysis of the effects of convergence on earnings. Donohue, Heckman and Todd (2002) study the sources of convergence in the South over the 1911-1960 time period.

4 Empirical Results

Section three described alternative approaches to estimating the cognitive achievement production function and the identifying assumptions that justify their application. A benefit of the rich longitudinal data we use is that they enable estimation of the general specifications, which are robust in the presence of unobserved endowments and endogeneity of input choices. We can also carry out formal tests of many of assumptions of more restrictive specifications. For this purpose, we use two types of specification tests. One is a general test of the null hypothesis that the model is correctly specified against the composite alternative hypothesis that it is misspecified. The other type of tests we use are standard Hausman and Wu tests (*e.g.*, Hausman, 1978, Wu, 1973, Godfrey, 1990) that compare the null hypothesis model against a specific alternative model. In what follows, we implement alternative specifications, moving from the most restrictive to the least restrictive, and, when possible, testing their identifying assumptions.

As discussed in section three, the contemporaneous specification places strong restrictions on the production technology but is less demanding than other specifications in terms of data requirements. Under the null that the contemporaneous model is correctly specified, test scores are a function only of contemporaneous input measures. A straightforward test of the contemporaneous specification that is implementable when historical data on inputs are available, is to include the historical input measures and check whether their associated coefficients are significantly different from zero.

The last two columns of Table 2a, labeled "OLS", present results for the cumulative specification estimated by ols for the PIAT Math Test Score Measure. Tables 2b and 2c present analogous results for the PIAT Reading Score and for the Composite Score (Math + Reading). The OLS specification includes the contemporaneous measure of the home input (Current Home) and the four historical measures that were defined in section three. The specification is given by

$$\begin{aligned} T_a &= \alpha_1 CurrentHome_a + \alpha_2 LagHome_a I(age \geq 8) + \alpha_3 LagLagAvg_a I(age \geq 10) \quad (9) \\ &\quad + \alpha_{4a} Home_{3-5} + \alpha_{5a} Home_{0-2} + X_a \beta + \varepsilon_a, \end{aligned}$$

where X represents other variables included in the specification.[35] The other variables (not shown in the table) are indicators for the child's age in years, birthweight, indicator for first and second born, indicators for mother not working when child was 0-1 and when child was

[35] In the specification, the home score at ages 0-2 and 3-5 appears separately from the home score at later ages, because of differences in the questions that make up the score. See discussion on this point in footnote 21 in section 3.

0-3, child gender, indicator for mother's age at birth < 18, 18-19 and 20-29, age in months of the child and age in months squared, mother's years of schooling, mother's AFQT percentile score, and the number of times the child has taken the test. These variables reflect either direct inputs, such as mother's employment, or biological endowments, such as birthweight.[36] We do not include race/ethnicity and family income in the production function specification, because these variables are neither direct inputs nor endowments.

The issue of whether or not to include proxy variables in the presence of omitted inputs is a difficult one, because the use of proxy variables that are correlated with included inputs can confound the interpretation of estimated model coeffients. Consider, for example, a model that relates achievement to home inputs. To compensate for missing data on home inputs a researcher might include family income. However, holding family income constant, an increase in an observed purchased input (such as books) implies lower expenditures on other potential unobserved inputs (such as paid tutors). Thus, when income is held constant, the estimated effect of the observed purchased input is misstated, because its effect is confounded with the effect of the change in the unobserved inputs. When proxy variables are correlated with included inputs, it is unclear whether including them in the specification reduces or increases bias.

For the ols specification, the estimated coefficients associated with lagged home input measures are significantly different from zero at conventional levels, which is evidence against a contemporaneous specification. The effect of lagged home inputs on test scores is similar in magnitude to that of the contemporaneous input measure. The estimated school input coefficients are of the expected sign, but only the effect of teacher salary is reasonably precisely estimated.[37]

If we estimate a specification that omits the lagged input measures (not shown in the table), for PIAT-Math we obtain a coefficient on Current Home of 0.052 with a standard error of 0.007. Thus, omitting historical measures leads to an overstatement of the impact of a unit increase in Current Home input. However, by neglecting the influence of historical measures, it also understates the impact of a unit increase in the home score sustained over

[36] The specification does not include presence of father in the household, because in many cases that variable is missing. When we did include it and estimated the equation for the subsample for which it is available, its coefficient was imprecisely estimated.

[37] A school input that we cannot measure is the curriculum content within the classroom. A proxy for curriculum could be the grade level the child is currently attending. However, to the extent that grade progression depends on prior achievement, grade level would reflect all past inputs and would be inappropriate to include. Also, if grade progression were automatic, age effects included in our specification would capture grade-specific curriculum content.

an extended time period. For example, the ols specification that includes historical measures implies that a unit increase in the home score at ages 6, 7 and 8 increases the age 9 PIAT Math test score by 0.079, which is fifty percent larger than the implied impact obtained from the contemporaneous specification. Results are similar for the PIAT-Reading Score and the Total Score.[38]

A similar specification test can be used to examine the support for the value-added model, which augments the contemporaneous specification with a lagged test score measure. The key assumption of the value-added model is that the lagged test score is a sufficient statistic for historical inputs and, in the versions of the model that do not incorporate endowments, the lagged test score is also taken to be a sufficient statistic for endowments. To test the first assumption, we include lagged input measures in the value-added specification, which should have no additional explanatory power under the sufficiency assumption. The estimates shown in the columns labeled "Value-Added" show that for Math and Total, the lagged home input measure is statistically significantly different from zero, and for Reading the lagged home measure and the home score at ages 3 to 5 are statistically significant. We interpret these results as evidence against the sufficiency assumption.[39] For the value-added model and for the total score, the school input variables are statistically significant and of the expected sign at a 10% level. However, the magnitude of their effects is not very large. The estimates imply, for example, that a change in the average pupil-teacher ratio by five fewer students would lead to a 0.7 increase in the total test score (the average total test score for the sample is 80 for whites, 71 for blacks and 73 for Hispanics). A $10,000 increase in teacher salary (in 1989 dollars) would lead to a 0.9 increase in the total test score.

In addition to the tests described above, we also test the null hypothesis that a particular model is correctly specified against a more general alternative specification using Hausman-Wu tests.[40] Table 3 describes the specifications tested and shows the p-values from each of the tests, where all of the specifications are derived from equation (9). The cumulative model that allows for child- or mother-specific unobserved endowments (fixed effects) nests the cumulative model with endowments that are orthogonal to included inputs (random effects)

[38] In the specification for PIAT-Reading, the coefficient on Current Home is 0.049 (0.007). For the Total Score, it is 0.102 (0.012). The ols specification that includes historical measures implies that a unit increase in the home score over ages 6,7, and 8 increases the Reading score by 0.11 and the Total score by 0.188, about double the impact implied by the contemporaneous specification.

[39] The estimated coefficient on the baseline test score is 0.6 for math, 0.6 for reading and 0.7 for Total, which does not support a first difference version of the value-added specification.

[40] A Hausman-Wu test requires that under the null, both the null hypothesis estimator and the alternative estimator are consistent, while under the alternative only the alternative estimator is consistent.

and also the versions of the model without endowments. Under the null that endowments are uncorrelated with inputs, the ols estimator applied to (4) is consistent, but under the alternative it is inconsistent. We base our test on a comparison of estimated coefficients for the model with endowments assumed to be orthogonal (a random effects model) and a fixed effects model.[41]

The column labeled "Mother F.E." gives the estimated coefficients for the cumulative specification with mother-specific endowment effects that may be correlated with input choices. A Hausman-Wu test comparing coefficient estimates from a random effect specification to estimates from the fixed effect specification rejects the random effects model with a p-value less than 0.01 for the PIAT-math test score measure, but not for the PIAT-reading test or for the composite test score. Thus, for the math test score, unobserved mother-specific endowments are not orthogonal to included inputs.

As described in section three, a cumulative model with child-specific endowment effects nests a model with mother-specific endowment effects. A Hausman-Wu test comparing the child fixed effect specification against a mother fixed effect specification rejects the latter specification (p-value < 0.01) for all three test scores, which, together with the previous result, indicates that within child heterogeneity is an important feature of the data.

As seen in Tables 2a,b,c, allowing for child fixed effects has important consequences for the magnitudes of the estimated coefficients. A comparison of the "Mother F.E." column and the "Child F.E." column shows that the effect of home inputs (current and lagged) declines, particularly for the PIAT-Math score, when one allows for child-specific endowments. However, all of the home input variables are statistically significantly different from zero. Under both the mother fixed effect and the child fixed effect specifications, the school input measures are for the most part insignificantly different from zero and sometimes not of the expected sign.

The within-child estimator allows input choices to be correlated with a child's fixed endowment but assumes that, conditional on endowment, input choices do not respond to earlier achievement realizations. It is plausible, however, that parents might adjust their input choices in response to their child's earlier achievement outcomes. For example, if a

[41] The Hausman test-statistic is given by

$$N(\hat{\beta}_{H_A} - \hat{\beta}_{H_0})'(\hat{V}_{H_A} - \hat{V}_{H_0})(\hat{\beta}_{H_A} - \hat{\beta}_{H_0})\tilde{}\chi^2(k),$$

where k is the dimensionality of $\hat{\beta}_{H_A}$ and $\hat{\beta}_{H_0}$, N is the sample size, and $\hat{V}_{H_A}$ and $\hat{V}_{H_0}$ are the components of the variance-covariance matrix associated with $\hat{\beta}_{H_A}$ and $\hat{\beta}_{H_0}$. The test statistic takes this form when the estimator under the null is efficient. See, e.g., Godfrey (1990).

child had a poor reading teacher that led to lower achievement, parents may buy more books for the child or spend more time reading with the child (activities that increase the home score).

To permit home and school input choices to respond to earlier achievement realizations as well as to unobserved endowments, we implement instrumental variable within-child estimators of the kind that were described in section two. The variables included in the instrument set are birthweight, birth order, whether the mother worked when the child was age 0-1 and when child was age 0-3, mother's AFQT, mother's age, mother's schooling, gender of child, home score at age 3-5, home score at age 0-2, spouses earnings, birthweight of first born child, mother's schooling when first child was born, whether there are children in the household < 1 year old, number of children in the age ranges 1-2, 3-5, 6-13 and 14+, race indicators (white, black), indicators for age of the child in years, the actual age of the child in months and the age squared, difference in age in years indicators, difference in age in months, difference in age in months squared, and the difference in mother's schooling.[42] Table 4 examines the correlation of the instruments with the regressors (the home and school input measures). For each regressor, the instruments are jointly significant (p-value < 0.001).

The column labeled "IV Child Diff" presents the estimated coefficients associated with the home and school input variables. A comparison of the child fixed effect specification that assumes exogenous input choices, conditional on endowments, with the IV model rejects exogeneity of the input choices for all the test score measures at a 10% level. On the whole, the specification test results provide evidence that child-specific endowment effects are important, that input choices are correlated with endowments and that input choices are correlated with the unobserved components of achievement realizations (conditional on endowments). In light of these findings, only the child fixed effect instrumental variables specification would be expected to yield consistent estimates.

As seen in Tables 2a-c, the home input variables (current and lagged) are statistically significant at conventional levels for nearly every specification and for both reading and math scores. For the results based on the IV specifications, the coefficients associated with lagged input tend to be more precisely estimated than the coefficients associated with the current home input measures. All the estimated coefficients on the home input measures are of the expected sign. The school input variables are for the most part statistically insignificant in the fixed effect specifications, with the exception of the pupil-teacher ratio (significant at the 10% level) in the equation for the PIAT-reading test score.[43] In all the within specifications (mother and child), omitted inputs for which race or "permanent" income would have served

[42] The instrument set also includes indicator variables for whether any of these variables is missing.

[43] The estimated coefficients on the home input variables are not much affected by whether school input

as a proxy are accounted for, because they are non-time-varying for a given child and do not vary across siblings.[44] We also estimated within child IV specifications that included spouse's earnings as a potential proxy variable for omitted inputs. Although statistically significant for the PIAT-reading test, including spouse's earnings had only a minor effect on the home input coefficients.

In addition to the results reported in the tables, we also estimated specifications where the pupil-teacher ratio was measured at the county level rather than the state level.[45] The county level measure was usually insignificant for all the specifications. Our finding that the pupil-teacher ratio is statistically significant only when measured at the state level is consistent with other findings reported in the literature that have compared estimated effects of school quality on earnings, where school quality is measured at different levels of aggregation. For example, Card and Krueger (1996) found significant effects of school quality measures on earnings, where quality is measured at the state average level. Betts (1995) compares the estimated effect of pupil-teacher ratios and teacher experience on earnings under different levels of aggregation of the quality measures and finds that the quality measures are only significant when measured at the state level.[46]

5 Accounting for Sources of Racial Test Score Gaps

Using the production function estimates from the last section, we examine the extent to which differences in inputs can account for racial/ethnic disparities in test scores. Because the coefficients associated with the school input variables tended to be imprecisely measured in all of the fixed effect specifications, we focus attention on the contribution of home inputs.

The specification test results of the previous section rejected all of the more restrictive specifications and gave support only for the child differenced IV model, which allows for child specific endowment effects and endogenous inputs. In Figures 3(a) and 3(b), we examine the

variables are included. If school input variables are omitted, then the estimated coefficients for Math are 0.059 (0.054) for Current Home, 0.025 (0.012) for Lag Home, 0.022 (0.011) for Lag Lag Avg. For Reading, the coefficients are 0.009 (0.052), 0.076 (0.015) and 0.041 (0.013), and for the Total Score, they are 0.068 (0.080), 0.101 (0.022), and 0.064 (0.019).

[44] The race of the child is by definition the race of the mother in the NLSY dataset.

[45] County is the most detailed measure of location available for the NLSY79 respondents.

[46] Our state-level measure of quality differs in some ways from measures in the literature, where state often corresponds to a person's state of birth and it is assumed that the child is educated in their state of birth. In our case, the state measure gives the state level average quality at the time of the child's residence. If a child moves from one state to another, our average school input measures would change to reflect different levels of school inputs across states and to reflect the amount of time spent in each location.

fit of this model.[47] The figures compare actual test score gaps by age to the gap predicted under the model, for boys and girls age 6-13 by racial/ethnic group.[48] The estimated model captures key features of the data, such as the magnitude of the gap for each of the groups and its widening with age that is most pronounced for black boys. Because the estimated production function coefficients do not vary by race/ethnicity nor by sex, the widening gap in the predicted test scores arises from race/sex differences in inputs. As noted in section two (Table 1 and Figure 2), there is a relatively constant disparity in home inputs between whites and minorities of different ages. A constant disparity in home inputs produces a widening gap in test scores when lagged inputs matter in producing current achievement, as the estimates in Table 2a-2c show they do.

Table 5 examines how the predicted math and reading test score gaps vary if we set the levels of home inputs at the average levels observed for white children. To examine sensitivity to alternative specifications, we report estimates for the ols model, the mother and child fixed effects models and the child differenced IV model. The column labeled "Pred. Gap with White Home Inputs" gives the predicted gap in average test scores if home inputs are equalized at the white average level but school inputs are kept at their race-specific values. The preferred IV estimates indicate that if black children received the white average levels of home inputs, the math test score gap would be reduced by 27% and the reading test score gap by 26%.[49] For Hispanic children, equalizing home inputs to white levels would reduce the gap by about 29% for math and 33% for reading. Estimates are for the most part of similar magnitudes across the different model specifications, except that the mother fixed effect specification yields a much larger reduction in the reading gap and the child fixed effect specification a smaller reduction in the math gap.

Our estimates imply that policies that would equalize home inputs of whites and blacks would close a significant proportion of the test score gap. However, a comparison of the efficiency of such policies would require information both about their ability to modify behavior and about the costs of implementation. In addition, the existence of such programs would potentially alter the levels of other inputs, which might augment or diminish their effectiveness. A full assessment of such policies would require a complete analysis of how

[47] The estimated model coefficients are reported in the second columns of Tables 2a-2c.

[48] The figures are based on the sample used to estimate the cognitive achievement production function. In forming these predictions, the fixed effects for each child are kept at their original estimated values. The fixed effects capture endowments and the effects of inputs prior to age 6, such as the home score at ages 0-2 and 3-5, birthweight, and mother's age at birth.

[49] In the fixed effect specifications, inputs prior to age 6 are included in the fixed effects. Therefore, when we equalize home inputs in Table 5, the home inputs refer only to inputs applied at age 6 or later.

families make decisions about what inputs to provide for their children.[50]

6 Conclusions

This paper considered ways of estimating the cognitive achievement production function that are consistent with theoretical notions that achievement is a cumulative process depending potentially on the entire history of family and school inputs as well as on unobserved endowments. Using rich longitudinal data, we implemented alternative specifications of the production function and tested assumptions that underlie commonly adopted estimating equations. We did not find support for restrictive specifications, such as the contemporaneous specification and the value-added model. The specification test results showed the importance of both contemporaneous and lagged inputs in the production of current achievement, of allowing for unobserved child-specific endowments and of allowing for endogeneity of inputs with respect to time varying components of children's achievement. The child differenced instrumental variables specification is consistent in the presence of these features of the data.

Across almost all the specifications considered, we found that home inputs (contemporaneous and lagged) are substantively significant determinants of child test scores. The magnitude of lagged home input effects is usually similar to that of current inputs. The coefficients associated with school inputs (pupil-teacher ratio and teacher salaries) were only found to be significant determinants of test scores in specifications that did not allow for fixed effects.

We used the production function parameter estimates to examine the sources of racial/ethnic test score gaps. Our estimated production function captures key features of the data, such as the widening of the minority-white test score gap with age. A striking feature of the results is that the predicted test score gaps (based on the estimated child differenced IV model) captures a significant portion of the pronounced widening of the gap for black boys, even though none of the model parameter estimates varies by race/ethnicity.

The contribution of home inputs to the test score gap was estimated to be similar for math and reading and for blacks and Hispanics. The results showed that equalizing home inputs would close about 25% of the black-white test score gap and about 30% of the Hispanic-white gap. Thus, our findings suggest that home input differences can account for a significant component of the gap. Moreover, the cumulated effect of a persistent disparity in home inputs can produce a widening of the test score gap that is observed for blacks.

[50] For example, see recent efforts by Mroz and Van der Klaauw (2003).

Our findings do not imply that the most efficient way to close the gap is to invest in ways of augmenting home inputs. What is required to make such determination is knowledge of the relative costs of alternative policies and of how schools and parents make input decisions, to account for the possibility that changing the level of a single input affects decisions about other inputs.

Finally, in all the specifications estimated, the home input variables tended to be more precisely estimated than the school inputs. A likely explanation for this pattern is that the home input variables are measured at the child-specific level, whereas the school variables are measured much more crudely at the state or county levels. To precisely estimate the effects of school inputs along with home inputs, it be desirable to have school inputs measured at the same level of aggregation as the home input measures (i.e. classroom level).

References

[1] Altonji, J. G. and Dunn, T. A. (1996). "Using siblings to estimate the effect of school quality on wages.'*Review of Economics and Statistics*, vol. 78, no. 4, (November), 665-71.

[2] Baharudin, R. and Luster, T. (1998). "Factors related to the quality of the home environment." *Journal of Family Issues*, vol. 19, no. 4, (July), 375-403.

[3] Baydar, N. and Brooks-Gunn, J. (1991). "Effects of maternal employment and child-care arrangements on preschoolers cognitive and behavioral outcomes: evidence from the Children of the NLSY." *Developmental Psychology*, vol. 27, 932-45.

[4] Ben-Porath, Y. (1967). 'The production of human capital and the life-cycle of Earnings," *Journal of Political Economy*, vol. 75, no. 4, part I, August, 352-365.

[5] Belsky, J. and Eggebeen, D. (1991). "Early and extensive maternal employment and young children's socioemotional development: Children of the National Longitudinal Survey of Youth." *Journal of Marriage and the Family*, vol. 53, 1083-1110.

[6] Betts, Julian (1995): "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth, " *Review of Economics and Statistics,* Vol. 77, 231-249.

[7] Blau, F. D. and Grossberg, A. J. (1992). 'Maternal labor supply and children's cognitive development.' *Review of Economics and Statistics,* vol. 74 (3), 474-81.

[8] Boardman, Anthony E. and Murnane, R. J. (1979), 'Using Panel Data to Improve Estimates of the Determinants of Educational Achievement," *Sociology of Education*, Vol. 52 (April): 113-121.

[9] Boozer, Michael A., Krueger, Alan B., and Shari Wolken (1992): "Race and School Quality Since Brown v. Board of Education." *Brookings Papers on Economic Activity-Microeconomics*, 269-326.

[10] Burtless, Gary (1996). 'Introduction and Summary.' Burtless, Gary, ed. *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Washington, DC: Brookings Institution.

[11] Caldwell and Bradley (1984). *Home Observation for Measurement of the Environment.* Little Rock: University of Arkansas at Little Rock, Center for Child Development and Education.

[12] Cameron, Stephen and Heckman, James J. (1993). 'The nonequivalence of high school equivalents.' *Journal of Labor Economics*, vol. 11, pp. 1-47.

[13] Caniero, Pedro, Heckman, James J. and Dimitriy V. Masterov (2002): "Labor Market Discrimination and Racial Differences in Premarket Factors," manuscript, University of Chicago.

[14] Caniero, Pedro and James J. Heckman (2003): "Human Capital Policy" in *Inequality in America: What Role for Human Capital Policies*, by Alan B. Krueger and James J. Heckman, ed. Benjamin Friedman, Cambridge, MA: MIT Press.

[15] Card, David and Alan B. Krueger (1996): "Labor Market Effects of School Quality: Theory and Evidence" in *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, ed. Gary Burtless, Washington, DC: Brookings Institution.

[16] Card, David and Alan B. Krueger (1992): "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics,* 107 (February), 151-200.

[17] Case, R. (1992). *Exploring the conceptual underpinnings of children's thought and knowledge*, Hillsdale, NJ: Erlbaum.

[18] Cook, Michael D. and William N. Evans (2000): "Families or Schools? Explaining the Convergence in White and Black Academic Performance," *Journal of Labor Economics,* Vol. 18, no.4, 729-753.

[19] Crane, Jonathan (1996): "Effects of home environment, SES, and maternal test scores on mathematics achievement." *The Journal of Educational Research*, vol. 89, no. 5, 305-314.

[20] Currie, Janet and Thomas, Duncan (1999). "Early test scores, socioeconomic status, and future outcomes." *NBER* working paper #6943.

[21] Desai, S., Chase-Lansdale, P. L. and Michael, R. T. (1989): "Mother or market? effects of maternal employment on the intellectual ability of 4-year old children." *Demography*, vol. 26, 545-61.

[22] Donohue, John, Heckman, James J. and Petra E. Todd (2002): "The Schooling of Southern Blacks: The Roles of Legal Activism and Private Philanthropy, 1910-1960," *Quarterly Journal of Economics,* Feb, 225-268.

[23] Fryer, Roland G. Jr.and Steven D. Levitt (2004): "Understanding the Black-White Test Score Gap in the First Two Years of School," *Review of Economics and Statistics*, Vol. 86, No. 2, May 2004, pp. 447-64.

[24] Fryer, Roland G. Jr. and Steven D. Levitt (2005): "The Black-White Test Score Gap Through the Third Grade," NBER working paper #11049.

[25] Fuchs, Victor R. and Diane M. Reklis (1994): "Mathematical Achievement in Eighth Grade: Interstate and Racial Differences," NBER working paper #4784.

[26] Godfrey, L. G. (1990). *Misspecification Tests in Econometrics*, Cambridge University Press: Cambridge.

[27] Hanushek, E. A. (1996). 'School resources and student performance.' in Burtless, Gary, ed. *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Washington, DC: Brookings Institution.

[28] Hanushek, E. A. (1998). Conclusions and controversies about the effectiveness of school resources.' *Economic Policy Review,* New York: Federal Reserve Bank of New York, vol. 4, no. 1, (March), pp. 11-28.

[29] Hausman, J. J. (1983). 'Specification tests in econometrics.' *Econometrica*, vol. 46, pp. 1251-71.

[30] Hedges, Larry V. and Amy Nowell (1999): "Changes in the Black-White Gap in Achievement Test Scpres," in *Sociology of Education,* Vol. 72 (April):111-115.

[31] Hedges, L. and Nowell, A. (1998): "The Black-White Test Score Convergence since 1965," in *The Black-White Test Score Gap*, Washington, DC: Brookings Institution, 149-181.

[32] Hutchinson, D., H. Prosser and P. Wedge (1979): "The Prediction of Educational Failure," *Educational Studies*, 5, #1.

[33] Jencks, Christopher and Meredith Phillips (1998): "The Black-White Test Score Gap: An Introduction," in *The Black-White Test Score Gap,* Washington, DC: Brookings Institution.

[34] Keane, M. P. and Wolpin, K. I. (1997). 'Career decisions of young men.' *Journal of Political Economy*, vol. 105, (June), 473-522.

[35] Krueger, Alan B. (1998). 'Reassessing the view that american schools are broken.' *Economic Policy Review*, publication of the Federal Reserve Bank of New York, March, vol. 4, no. 1., p. 29-46.

[36] Krueger, Alan B. (2003): "Economic Considerations and Class Size" in *The Economic Journal,* Vol. 113, No. 485, F34-63.

[37] Leibowitz, Arleen (1974): 'Home investments in children.' *Journal of Political Economy*, vol. 82(2), Pt. I, (March/April), pp. S111-31.

[38] Ludwig, Jens (2003): "The great unknown: does the black-white test score gap narrow or widen through the school years? It depends on how you measure," in *Education Next*, June.

[39] Murnane, R. J., Maynard, R. A. and Ohls, J. C. (1981). 'Home resources and children's achievement.' *The Review of Economics and Statistics*, vol. 63, no. 3, pp. 369-377.

[40] Murnane, R. J. Willet, J. B. and Levy, F. (1995). 'The growing importance of cognitive skills in wage determination.' *Review of Economics and Statistics*, LXXVII, pp. 251-266.

[41] Neal, D. and Johnson, W. (1996). 'The role of pre-market factors in black-white wage differences.' *Journal of Political Economy*, vol. 104, no. 5, (October), pp.869-895.

[42] Parcel, T. L. and Menaghan, E. (1994). 'Early parental work, family social capital and early childhood outcomes.' *American Journal of Sociology*, vol. 99, no. 4, pp. 972-1009.

[43] Philips, M., Crouse, J. and J. Ralph (1998): "Does the Black-White Test Score Gap Widen after Children Enter School" in *The Black-White Test Score Gap*, Washington, DC: Brookings Institution, 229-272.

[44] Robertson, Donald and James Symons (1990): "The Occupational Choice of British Children," in *The Economic Journal,* 100, 828-841.

[45] Rosenzweig, M. (1986): "Birth spacing and Sibling Inequality: Asymetric Information within the Family," in *IER*, Vol. 26, No. 3, 55-76.

[46] Rosenzweig, M. and Schultz, T. P. (1982). "Market opportunities, genetic endowments and intrafamily resource distribution: child survival in rural India," *American Economic Review*, vol. 72, no. 4, (September), 803-815.

[47] Rosenzweig, M. and Wolpin, K. I. (1988). "Heterogeneity, intrafamily distribution and child health.' *Journal of Human Resources*, vol. 23, no. 4, (Fall), 437-461.

[48] Rosenzweig, M. and Wolpin, K. I. (1994): “Are there increasing returns to the intergenerational production of human capital? Maternal schooling and child intellectual achievement,” *The Journal of Human Resources*, vol. 29, no. 2, (Spring), 670-693.

[49] Rosenzweig, M. and Wolpin, K. I. (1995). ‘Sisters, siblings, and mothers: the effects of teen-age childbearing on birth Outcomes.’ *Econometrica*, vol. 63, no. 2, (March), 303-326.

[50] Summers, Anita and Barbara Wolfe (1977): “Do Schools Make a Difference?" *American Economic Review*, Vol. 67, No. 4, 639-652.

[51] Todd, Petra and Kenneth I. Wolpin (2003): “On the Specification and Estimation of the Production Function for Cognitive Achievement," in *The Economic Journal,* Vol. 113, No. 485, F3-33.

[52] Vandell, D. and Ramanan, J. (1992). ‘Effects of early and recent maternal employment on children from low income families,’ *Child Development*, vol. 63, 938-949.

[53] Wu, D. M. (1973). ‘Alternative tests of independence between stochastic regressors and disturbances.’ *Econometrica*, vol. 41, pp. 733-50.

TABLE 1
Descriptive Statistics: Means with Standard Errors in Parentheses

	White	Black	Hispanic
Piat Math (Age 6-13)	40.6	35.6	36.3
	(15.2)	(14.4)	(15.0)
Piat Reading (Age 6-13)	40.0	35.6	36.6
	(15.1)	(13.3)	(14.8)
Current Home Score	102.1	86.6	88.0
	(21.9)	(24.5)	(24.6)
Lag Home Score	102.0	86.7	88.2
	(21.9)	(24.5)	(24.5)
Lag-Lag Average Home Score	102.4	87.2	89.0
	(21.5)	(24.4)	(24.2)
Average Home Score Age 3-5	122.2	104.8	106.2
	(18.2)	(24.4)	(25.0)
Average Home Score Age 0-2	70.7	60.8	62.8
	(12.7)	(15.2)	(14.6)
Average (over all of the child's	18.5	18.1	20.0
school years) Pupil-Teacher Ratio	(2.5)	(2.3)	(3.5)
Average (over the child's school			
years) Teacher Salary (1989 $)	31,387	29,342	32,121
Child Age (in months)	114.9	117.3	115.9
	(27.0)	(27.2)	(27.1)
Birthweight (ounces)	120.3	111.1	117.2
	(20.2)	(21.7)	(20.7)
Percent Firstborn	50.4	43.4	43.9
Percent Second born	34.0	32.2	32.4
Percent Mother's not working			
Child Age 0-1	37.6	48.4	48.2
Child Age 0-3	20.3	28.1	27.7
Percent Mother's age at birth			
Less than 18	4.6	12.4	7.4
18-19	9.5	16.4	15.1
20-29	75.8	66.3	71.8
30+	10.1	4.9	5.6
Mother's Schooling	12.7	12.1	11.2
	(2.2)	(2.0)	(2.7)
Mother's AFQT Percentile Score	49.0	18.9	22.8
	(25.9)	(17.2)	(20.7)

TABLE 2a

Estimated Effects of Home and School Inputs under Alternative Specifications of the Educational Production Function

Dependent Variable: PIAT Math Score*

	IV Child Diff	Child F.E.	Mother F. E.	Value-Added	OLS
Home Inputs					
(1) Current Home	0.069	0.012	0.031	0.029	0.030
	(0.056)	(0.005)	(0.007)	(0.007)	(0.007)
(2) Lag Home	0.023	0.014	0.022	0.021	0.028
	(0.012)	(0.004)	(0.006)	(0.007)	(0.006)
(3) Lag Lag Avg Home	0.024	0.010	0.016	-0.004	0.015
	(0.011)	(0.004)	(0.006)	(0.005)	(0.006)
(4) Home Age 3-5	…	…	0.022	0.003	0.032
			(0.011)	(0.007)	(0.009)
(5) Home Age 0-2	…	…	0.008	0.014	0.024
			(0.015)	(0.009)	(0.012)
p-value: joint sig (1)-(3)	0.024	0.001	0.000	0.000	0.000
p-value: joint sig (2)-(3)	0.009	0.000	0.000	0.012	0.000
p-value: joint sig (4)-(5)	…	…	0.129	0.251	0.001
School Inputs					
(6) Avg Pupil-Teacher Ratio	-1.02	0.099	0.220	-0.098	-0.081
	(1.02)	(0.102)	(0.130)	(0.056)	(0.071)
(7) Avg Teacher Salary/1000	-2.516	-0.384	-0.320	0.068	0.095
	(2.183)	(0.137)	(0.1951)	(0.035)	(0.043)
p-value: joint sig (6)-(7)	0.147	0.016	0.070	0.121	0.085
Number of Observations	7269	13017	6852	4838	6667

* Additional variables in the ols specification are: indicators for child's age in years, birthweight, indicators for first and second born, indicators for mother not working when child was age 0-1 and when child was 0-3, child gender, indicator for mother's age at birth < 18, 18-19 and 20-29, age in months of the child and age in months squared, indicators for the mother being black and Hispanic, mother's years of schooling, mother's AFQT percentile score, and the number of times the child has taken the test. The fixed effect specifications include the same set of variables, except the ones that are constant (within mothers or within children).

TABLE 2b

Estimated Effects of Home and School Inputs under Alternative Specifications of the Educational Production Function

Dependent Variable: PIAT Reading Score*

	IV Child Diff	Child F.E.	Mother F. E.	Value-Added	OLS
Home Inputs					
(1) Current Home	0.019	0.025	0.030	0.019	0.024
	(0.054)	(0.005)	(0.007)	(0.007)	(0.007)
(2) Lag Home	0.078	0.030	0.290	0.018	0.036
	(0.015)	(0.004)	(0.006)	(0.007)	(0.006)
(3) Lag Lag Avg Home	0.042	0.024	0.018	-0.008	0.015
	(0.013)	(0.004)	(0.006)	(0.006)	(0.007)
(4) Home Age 3-5	…	…	0.033	0.015	0.028
			(0.012)	(0.007)	(0.009)
(5) Home Age 0-2	…	…	0.007	0.014	0.012
			(0.015)	(0.010)	(0.013)
p-value: joint sig (1)-(3)	0.000	0.000	0.000	0.000	0.000
p-value: joint sig (2)-(3)	0.000	0.000	0.000	0.019	0.000
p-value: joint sig (4)-(5)		….	0.017	0.014	0.002
School Inputs					
(6) Avg Pupil-Teacher Ratio	-1.78	-0.192	-0.109	-0.004	-0.103
	(1.06)	(0.111)	(0.135)	(0.051)	(0.070)
(7) Avg Teacher Salary/1000	-1.38	0.288	0.189	0.070	0.135
	(2.204)	(0.147)	(0.202)	(0.030)	(0.041)
p-value: joint sig (6)-(7)	0.062	0.053	0.478	0.034	0.004
Number of Observations	7269	13017	6859	4838	6667

* Additional variables included in the ols specification are: indicators for child's age in years, birthweight, indicators for first and second born, indicators for mother not working when child was age 0-1 and when child was 0-3, child gender, indicator for mother's age at birth < 18, 18-19 and 20-29, age in months of the child and age in months squared, indicators for the mother being black and Hispanic, mother's years of schooling, mother's AFQT percentile score, and the number of times the child has taken the test. The fixed effect specifications include the same set of variables, except the ones that are constant (within mothers or within children).

TABLE 2c
Estimated Effects of Home and School Inputs under Alternative Specifications of the Educational Production Function
Dependent Variable: PIAT Total Score (Math + Reading)*

	IV Child Diff	Child F.E.	Mother F. E.	Value-Added	OLS
Home Inputs					
(1) Current Home	0.088	0.037	0.058	0.0460	0.054
	(0.083)	(0.007)	(0.011)	(0.011)	(0.012)
(2) Lag	0.105	0.044	0.050	0.035	0.064
	(0.022)	(0.006)	(0.009)	(0.011)	(0.011)
(3) Lag Lag Avg	0.067	0.034	0.034	-0.017	0.031
	(0.019)	(0.007)	(0.009)	(0.009)	(0.012)
(4) Home Age 3-5	…	…	0.056	0.011	0.060
			(0.029)	(0.010)	(0.016)
(5) Home Age 0-2	…	…	0.015	0.024	0.036
			(0.024)	(0.015)	(0.023)
p-value: joint sig (1)-(3)	0.0000	0.0000	0.0000	0.000	0.000
p-value: joint sig (2)-(3)	0.0000	0.0000	0.0000	0.003	0.000
p-value: joint sig (4)-(5)	…	…	0.0105	0.105	0.000
School Inputs					
(6) Avg Pupil-Teacher Ratio	-2.81	-0.133	0.113	-0.099	-0.184
	(1.61)	(0.162)	(0.220)	(0.085)	(0.126)
(7) Avg Teacher Exp/1000	-3.90	0.169	-0.134	0.121	0.230
	(0.855)	(0.215)	(0.321)	(0.052)	(0.076)
p-value: joint sig (6)-(7)	0.022	0.461	0.809	0.068	0.010
Number of observations	7269	13017	6852	4838	6667

* Additional variables included in the ols specification are: indicators for the child's age in years, birthweight, indicators for first and second born, indicators for mother not working when child was age 0-1 and when child was 0-3, child gender, indicator for mother's age at birth < 18, 18-19 and 20-29, age in months of the child and age in months squared, indicators for the mother being black and Hispanic, mother's years of schooling, mother's AFQT percentile score, and the number of times the child has taken the test. The fixed effect specifications include the same set of variables, except the ones that are constant (within mothers or within children).

TABLE 3
P-values from Hausman-Wu Specification Tests

Specification Test	PIAT-Math	PIAT-Reading	PIAT – Total
	Test Score Measure		
Test of Mother Fixed Effect model (alt) against Mother Random Effect model (null)	0.005	0.820	0.388
Test of Child Fixed Effect Model (alt) against Child Random Effect model (null)	0.002	0.017	0.002
Test of Child Fixed Effect Model (alt) against Mother Fixed Effect model (null)	0.013	0.002	0.000
Test of IV Child Differenced (alt) against Child Differenced (without IV) (null)	0.077	0.039	0.002

TABLE 4
R-squared values from First Stage Regressions of Within-Child Differenced Input Variables on Instrument Set

	R-squared	p-value from joint test of significance
Differenced		
Current Home	0.0153	0.000
Lag Home	0.0292	0.000
Lag Lag Avg	0.0372	0.000
Avg Pupil-Teacher Ratio	0.0237	0.000
Avg Teacher Salary	0.0318	0.000

See Section 4 for description of the set of instruments.

TABLE 5
Percent of Racial Test Score Gap Explained by Home and School Input Differences

Racial/ Ethnic Group	Age	Test	Avg. Score	Avg. White Score	Gap	Pred. Gap w/ White Home Inputs	% Gap Explaine d by Home
Black	OLS	Math	36.0	41.0	5.0	4.2	15%
		Reading	35.6	40.0	4.4	3.7	17
	Mother FE	Math	36.0	41.0	5.0	4.3	14
		Reading	35.6	40.0	4.4	1.4	69
	Child FE	Math	36.0	41.0	5.0	4.7	6
		Reading	35.6	40.0	4.4	3.7	17
	Child Diff IV	Math	36.0	41.0	5.0	3.7	27
		Reading	35.6	40.0	4.4	3.3	26
Hispanic	OLS	Math	36.6	41.0	4.4	3.7	16
		Reading	36.5	40.0	3.5	2.8	20
	Mother FE	Math	36.6	41.0	4.4	3.7	16
		Reading	36.5	40.0	3.5	0.4	89
	Child FE	Math	36.6	41.0	4.4	4.1	7
		Reading	36.5	40.0	3.5	2.8	21
	Child Diff IV	Math	36.6	41.0	4.4	3.2	29
		Reading	36.5	40.0	3.5	2.4	33

Table A.1
Comparison of Responses on Individual Home Input Score Items by Race
Age 0-2 Questions

Question	White	African American	Hispanic
Asked of Mother			
How often does child get out of the house?	81	72	63
How many children's books does your child have? (1 if > 3)	90	63	65
How often do you read stories to your child? (1 if at least 3 times per week)	74	41	42
How often do you take your child to the grocery store?	35	40	45
How many cuddly, soft or role-playing toys does your child have? (1 if >0)	100	99	99
How many push or pull tors does your child have?	99	93	95
Some parents spend time teaching children new skills while other parents believe children learn best on their own. Which describes your attitude? (=1 if always or usually spends time teaching children)	93	96	94
Interviewer Observations			
Mother provided toys or interesting activities for child?	76	50	63
Child's play environment is safe (no dangerous health or structural hazards within a todder's range)	94	89	90

* Number denotes the percent receiving highest score, where highest score is 1 and low score is 0.

Table A.2
Comparison of Responses on Individual Home Input Score Items by Race
Age 3-5 Questions

Question	White	African American	Hispanic
*Asked of Mother**			
How many children's books does your child have? (1 if > 9)	94	57	63
How often do you read stories to your child? (1 if at least 3 times a week)	70	40	44
How many magazines does your family get regularly? (1 if at least 3)	41	31	26
Does your child have the use of a CD player or tape recorder or record player and at least 5 children's records or tapes?	83	66	67
Do you help your child learn numbers?	78	63	70
Do you help your child learn the alphabet?	77	63	66
Do you help your child learn colors	78	62	70
Do you help your child learn shapes/sizes?	72	51	56
How often does a family member take the child on any kind of outing?	87	74	77
How often does family member arrange visit to museum within last year?	72	66	61
Interviewer Observations			
Child's play environment is safe (no dangerous health or structural hazards within a preschooler's range)	95	90	91
Is interior of the home dark or perceptually monotonous?	5	16	8
Are visible rooms reasonably clean?	94	90	92
Are visible rooms minimally cluttered?	84	82	84

* Number denotes the percent receiving highest score, where highest score is 1 and low score is 0.

Table A.3
Comparison of Responses on Individual Home Input Score Items by Race
Age 6-9 Questions

Question	White	African American	Hispanic
*Asked of Mother**			
How many books does your child have? (1 if > 9)	94	64	68
How often do you read to your child?	45	28	32
Is there a musical instrument that your child can use at home?	47	30	33
Does family get a daily newspaper?	51	41	40
How often does child read for enjoyment?	74	72	68
Does family encourage child to start and keep hobbies?	93	86	83
Does child get special lessons or belong to organizations that encourage activities (sports, arts)	61	41	39
How often does family member arrange visit to museum within last year?	80	70	70
How often does family member take child to musical or theatrical performance within the past year?	61	56	49
When family watches TV together, do you or child's father discuss TV programs with him/her	88	71	79
Interviewer Observations			
Is interior of the home dark or perceptually monotonous?	5	15	7
Are visible rooms reasonably clean?	93	90	92
Are visible rooms minimally cluttered?	84	83	84
Building has no potentially dangerous structural or health hazards within's a school-aged child's range	71	65	70

* Number denotes the percent receiving highest score, where highest score is 1 and low score is 0.

Table A.4
Comparison of Responses on Individual Home Input Score Items by Race
Age 10-13 Questions

Question	White	African American	Hispanic
*Asked of Mother**			
How many books does your child have? (1 if > 19)	75	38	43
Is there a musical instrument that your child can use at home?	55	32	35
Does family get a daily newspaper?	49	41	41
How often does child read for enjoyment?	66	64	65
Does family encourage child to start and keep hobbies?	95	89	88
Does child get special lessons or belong to organizations that encourage activities (sports, arts)?	69	56	50
How often does family member arrange visit to museum within last year?	80	72	69
How often does family member take child to musical or theatrical performance within the past year?	61	59	50
When family watches TV together, do you or child's father discuss TV programs with him/her	87	69	74
Interviewer Observations			
Is interior of the home dark or perceptually monotonous?	5	15	8
Are visible rooms reasonably clean?	93	91	93
Are visible rooms minimally cluttered?	85	83	83
Building has no potentially dangerous structural or health hazards within a school-aged child's range	70	65	66

* Number denotes the percent receiving highest score, where highest score is 1 and low score is 0.

Table B.1
Correlation between Most Recent Piat Score Age ≥12 and Earlier PIAT Scores

	Most Recent PIAT	$PIAT_{-1}$	$PIAT_{-1}$	$PIAT_{-1}$
Most Recent PIAT	1.0	-	-	-
$PIAT_{-1}$	0.78	1.0	-	-
$PIAT_{-2}$	0.72	0.78	1.0	-
$PIAT_{-3}$	0.58	0.65	0.73	1.0

Table B.2
Correlation between Most Recent PIAT Score and Highest Grade Completed at Indicated Age Ranges

Age	Correlation Coefficient	Number of observations
18+	0.33	917
19+	0.37	656
20+	0.35	477
21+	0.32	306
22+	0.38	201

Figure 1a: Comparison of PIAT−Reading Scores by Age by Race/Ethnicity

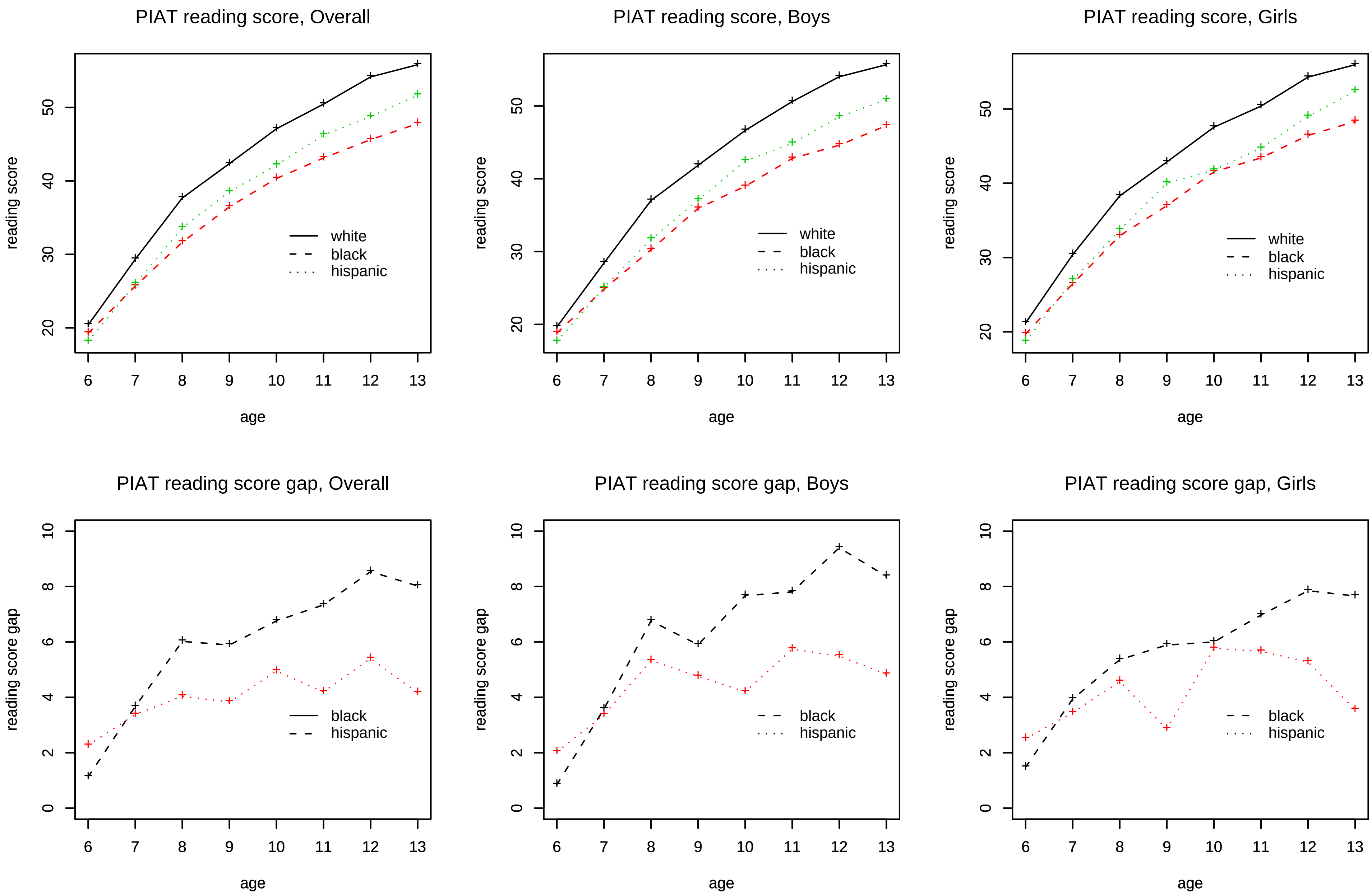

Figure 1b: Comparison of PIAT−Math Scores by Age by Race/Ethnicity

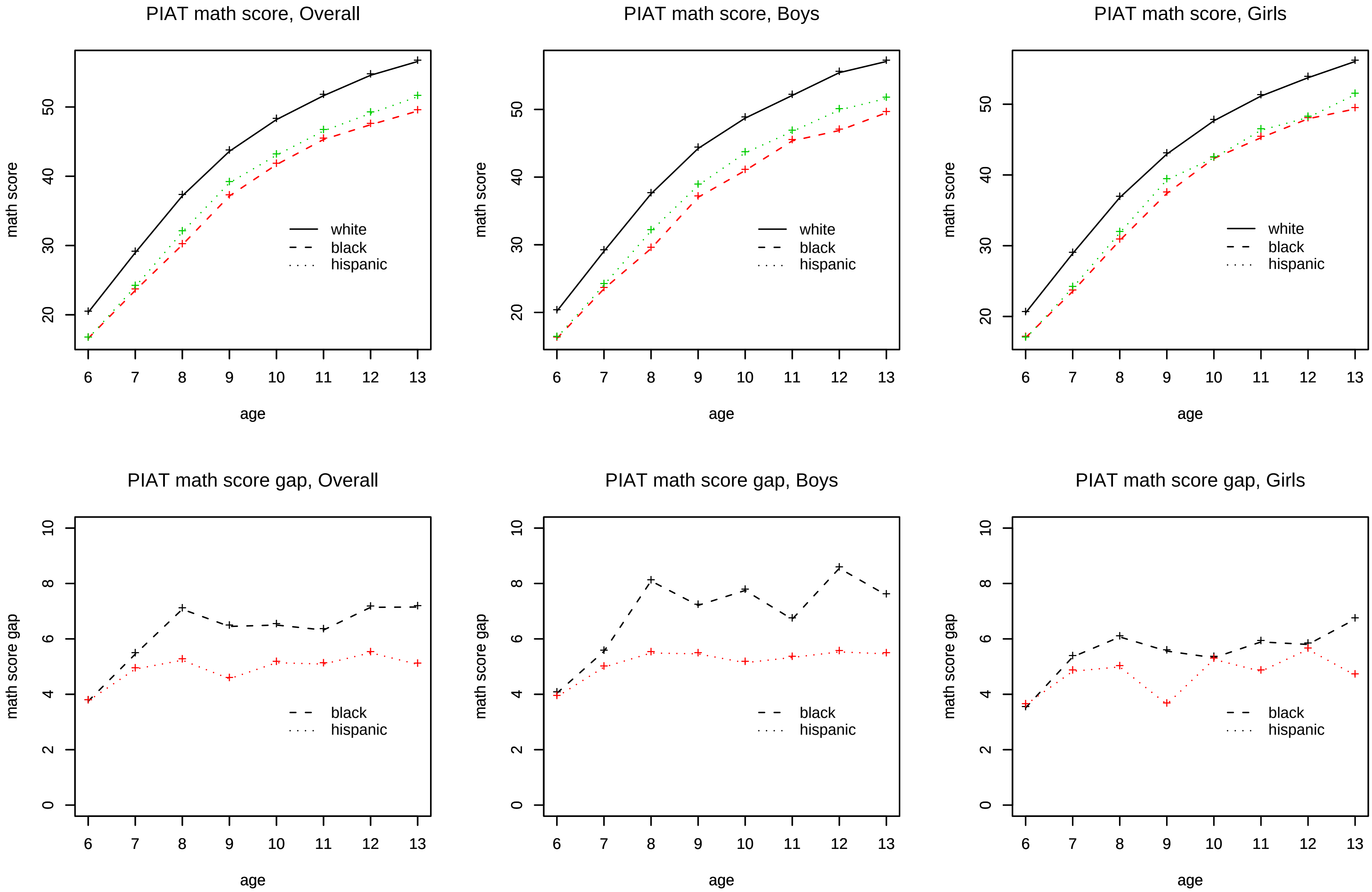

Figure 2: Comparison of Current Home Score by Age by Race/Ethnicity

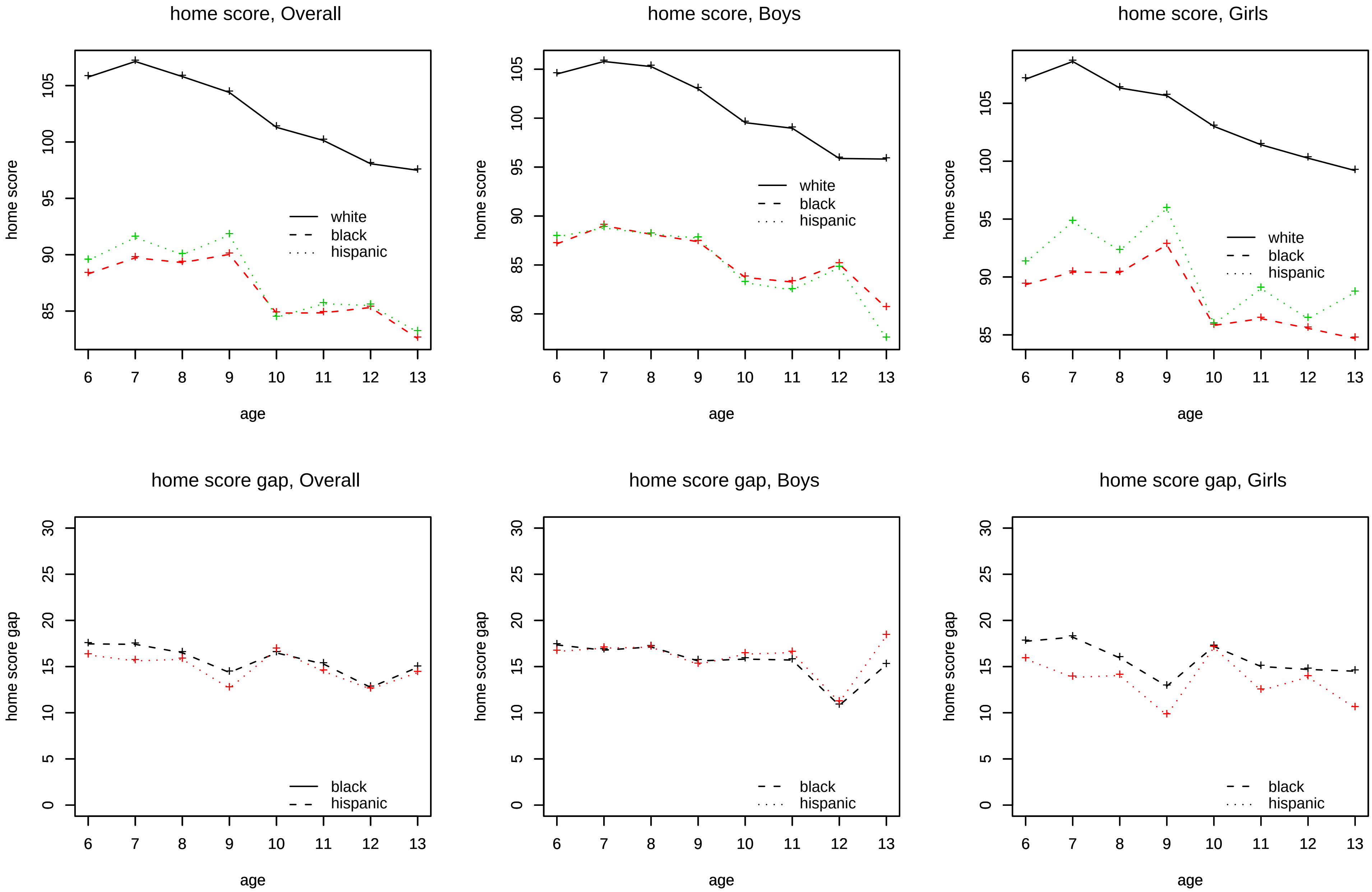

Figure 3(a): Actual and Predicted Math and Reading Test Score Gaps for Estimation Sample by Age and by Race/Ethnicity

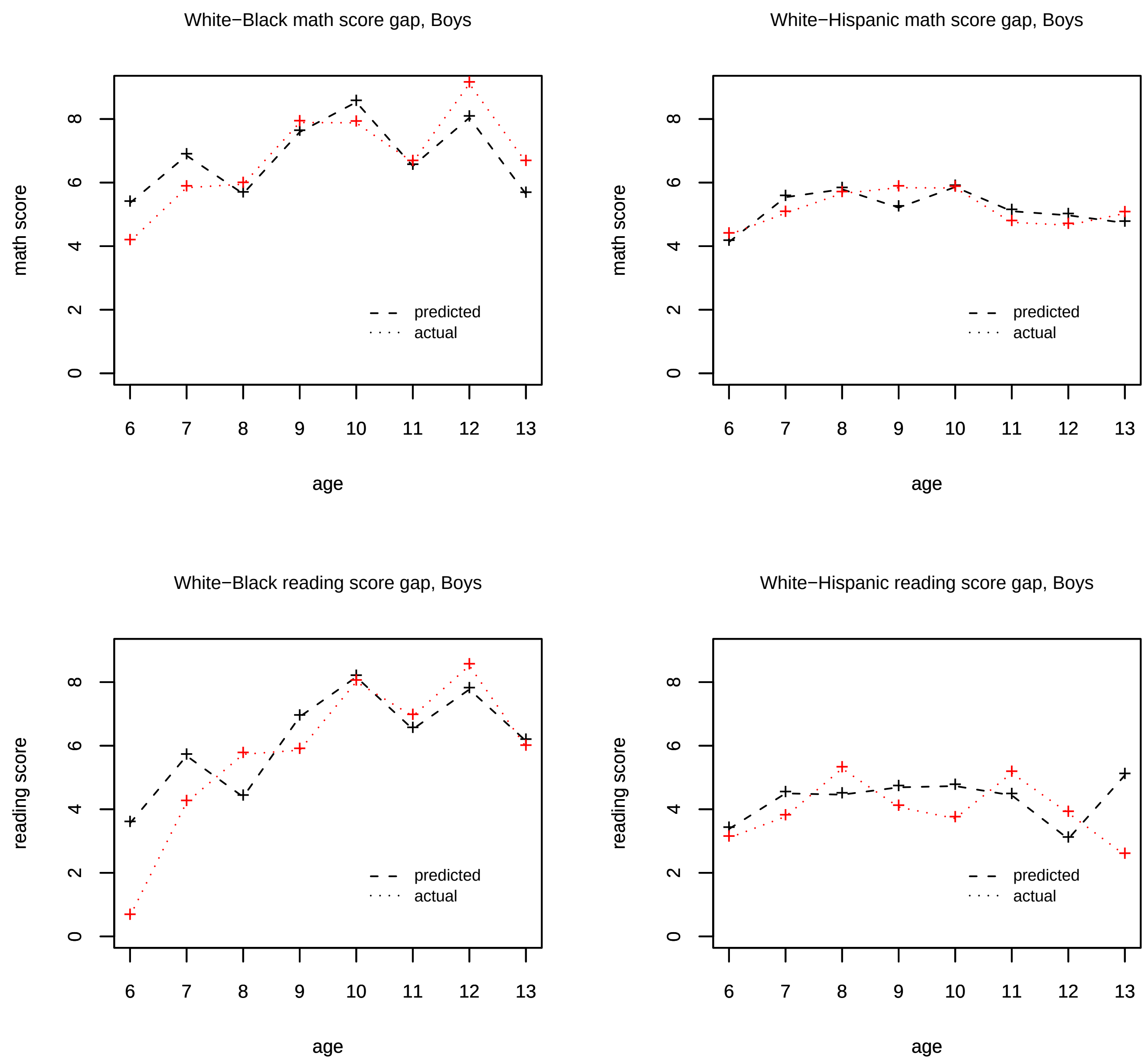

Figure 3(b): Actual and Predicted Math and Reading Test Score Gaps for Estimation Sample by Age and by Race/Ethnicity