Permanent Income and the Black-White Test Score Gap

Jesse Rothstein
Princeton University and NBER

Nathan Wozny*
Princeton University

June 2009

Abstract

Analysts often examine the black-white test score gap conditional on family income. Theoretically, we are interested in the gap conditional on permanent income, but typically only a current income measure is available. We develop methods for identifying the gap conditional on permanent income, using an auxiliary data set to estimate the reliability of current income. Current income explains only about half as much of the black-white test score gap as does permanent income, and the remaining gap in math achievement among families with the same permanent income is only 0.2 to 0.3 standard deviations in the CNLSY and ECLS samples. When we add permanent income to the controls used by Fryer and Levitt (2006), the unexplained gap in 3rd grade shrinks to less than 0.2 SDs, around half of what is found with their controls.

* Corresponding author. E-mail nwozny@princeton.edu. We thank Lars Lefgren and Richard Rothstein for helpful conversations, and Fanyin Zheng for excellent research assistance.

I. Introduction

The black-white test score gap has been extensively documented. On many different types of tests, administered to children of various ages over several decades, average scores for blacks are substantially lower than those for whites. Although the precise magnitude of the gap varies across samples, tests, and ages, gaps approaching one full standard deviation are not uncommon.

A central question in studies of the black-white test score gap has been whether it can be explained by differences in families' "objective" characteristics. If, for example, there were no black-white gap among families with the same income, we might hope that eventual convergence of black and white family incomes (Krueger, Rothstein, and Turner 2006; Neal 2006) would lead to the disappearance of the test score gap.

Ethnographic evidence (Kozol 1992, Lareau 2003) suggests that material circumstances can account for much of the black-white gap. But statistical evidence has not supported this view. Jencks and Phillips (1998) summarize the state of knowledge: "Income inequality between blacks and whites appears to play some role in the test score gap, but it is quite small" (p. 9); or, elsewhere, "the gap shrinks only a little when black and white families have the same amount of schooling, the same income, and the same wealth" (p. 2). Fryer and Levitt (2004, 2006) are the most successful at explaining the gap via differences in observable characteristics, but even they find that the gap that remains after controlling for a vector of demographic and behavioral variables is nearly 0.4 standard deviations by the end of the 3rd grade (down from a raw gap of 0.88 SDs in their sample). The robustness of the black-white gap to the inclusion of extensive controls has often been taken to support the view that the gap is largely attributable to differences in genes (Herrnstein and Murray 1996), culture (Moynihan 1965), or parenting styles (Brooks-Gunn et al. 1996) between blacks and whites. These factors are unlikely to be amenable to simple policy interventions.

Many studies of the conditional black-white gap control for the family's income, nearly always measured in the year that the child was tested. It has long been recognized that annual income is a poor proxy for a family's financial resources (Modigliani and Brumberg 1954; Friedman 1957), and in literatures where income is the explanatory variable of interest, researchers have long attempted to form better measures of permanent income. One common strategy has been to average incomes over several years, perhaps five (e.g., Solon 1992, Mayer 1997). More recent studies have suggested that even short-run averages are quite noisy measures of permanent income (Mazumder 2001, Haider and Solon 2006). Using a longer-term income average, Mazumder (2003) concludes that measurement error in 5-year-income averages leads to attenuation of the intergenerational elasticity of income by nearly 30%. But the insight that it is important to measure permanent income accurately has been slow to penetrate literatures where income is used only as a control variable, despite the well-known result that mismeasurement of one explanatory variable will bias the coefficients for all right-hand-side variables in OLS regressions. Researchers typically use annual income (Campbell et al. 2008) or short-run averages (Phillips et al. 1998; Blau and Grossberg 1992), or simply substitute other variables – like maternal education or socioeconomic status indices – that are thought to be effective proxies for permanent income (Fryer and Levitt 2004, 2006).[1]

In this paper, we argue that reliance on poor income measures has led past research to substantially understate the explanatory power of differences in family income for the black-white test score gap. Consider a regression that controls for annual income, where test scores are instead related to permanent income. Because annual income is a noisy proxy for the true variable, its coefficient is attenuated relative to what would be observed if permanent income were controlled. As mean permanent income is lower for blacks than for whites, this attenuation leads to overstatement of the black-white test score gap conditional on income. Using data from the National Longitudinal Study of Youth (NLSY) Child Supplement, we show that the bias is substantial: Only half of the black-white gap in the permanent income of the families of 10 and 11-year-old children is eliminated by controlling for annual income.

One reason that researchers do not control for permanent income is that the data requirements are onerous: One needs to observe child achievement and a long family income history, and the two are rarely available in the same data set. We develop methods that can be used to identify the black-white test score gap conditional on

[1] An exception is Blau (1999), who finds that average incomes over 12 years have only small effects on child development.

permanent income even when the test score data contain only annual income measures, using an auxiliary data set (like the NLSY) with more complete income histories.

We propose two estimators for the conditional gap. The first is a two-sample two-stage-least-squares estimator that uses annual income as an instrument for permanent income. The second uses the traditional errors-in-variables formula for OLS regression to recover the coefficient of interest. Each estimator can be applied even when the data set containing student test scores lacks a permanent income measure. We require only that the primary data set report test scores, race, and current income, and that the auxiliary data set have race and both current and permanent income measures. The former variables are available in nearly all educational data sets, so our estimators can be used to identify the black-white test score gap conditional on family income in all of the samples that are typically used to study the black-white gap. Our errors-in-variables estimator can even be applied to specifications including covariates that are available only in the primary data set, under reasonable additional assumptions.

In data from the Early Childhood Longitudinal Study (ECLS), we find that the black-white math score gap at the end of 5th grade, controlling for permanent income and for a very short list of family structure variables (primarily the mother's age at the child's birth), is only 0.24 standard deviations. By contrast, the gap without income controls is 0.64 standard deviations and the traditionally-estimated gap conditional on annual income is 0.44. Moreover, when we add permanent income to the short vector of controls used by Fryer and Levitt (2004, 2006) – an SES composite, birth weight, number of children's books in the home, mother's age at first birth, and a WIC recipient indicator – we find that the remaining black-white gap in 3rd grade falls to 0.18, just over half of the 0.34 that we obtain without income controls. Many of the other variables have no effect once income is adequately controlled, suggesting that their strong explanatory power in the Fryer and Levitt analysis primarily reflects their value as proxies for family financial resources.

Our approach and methods are applicable beyond the analysis of test score gaps. An identical strategy would be useful whenever one is willing to impose our key identifying assumption, that transitory variations in annual income are unrelated with outcomes conditional on permanent income. This assumption is implied by economic theory in many contexts – frequently, it follows immediately from the permanent income hypothesis – and even when it does not hold exactly it may be a tolerably close approximation. Thus, one might use our strategy to estimate the gap in homeownership or life expectancy between blacks and whites with the same permanent income, for example.

In Section II, we provide an overview of the black-white gap and of the role of income controls in studies of this gap. In Section III, we develop a simple econometric model of the role of income and race, and we describe our two approaches to identifying the black-white gap conditional on permanent income. In Section IV, we discuss the two data sets used in this paper, the NLSY-C and the ECLS. In Section V, we present simple analyses of the dynamics of family income in the NLSY data. Section VI presents results on black-white test score gaps. Section VII concludes.

II. Overview of the Problem

A. The raw black-white test score gap

The gap in mean scores between black and white children is large and persistent, nearly always above 0.5 standard deviations and more commonly in the 0.75 to 1 range.[2] It is robust across many different samples and different tests, and in particular is apparent on both aptitude and achievement tests. It tends to be somewhat larger for older than for younger children (Phillips et al. 1998, Fryer and Levitt 1996, 1998). In longitudinally consistent data, it is smaller for children born since 1970 than for those born before 1960, though there was little additional progress for children born between the early 1970s and late 1980s (Chay, Guryan, and Mazumder 2009).

B. The role of income in educational production

To understand why researchers often focus on the black-white gap in test scores conditional on income and other observable characteristics, it is useful to consider scores as an outcome of a household optimization process. We assume that a unitary household with income y must divide this income between educational investments, e, and

[2] See, e.g., the reviews by Neal (2006), Jencks and Phillips (1998), and Magnuson and Waldfogel (2008).

numeraire consumption, c: $y = c + e$. The child's educational achievement, s, is a function of innate aptitude, a, and investments: $s = f(a, e)$. The household's utility depends on consumption, the student's achievement, and preference parameter γ: $u=U(s, c; \gamma)$.[3] This preference parameter might capture variation in parents' altruism, in parenting styles, or in direct tastes for education.[4]

Thus, the household's problem is to choose e to maximize $U(f(a, e), y - e; \gamma)$. The implicit function theorem implies that we can write the chosen expenditure level as a function of ability, income, and preferences, $e^* = g(a, y, \gamma)$. The realized test score will then be $s^* = f(a, g(a, y, \gamma)) \equiv h(a, y, \gamma)$. Taking a first-order (linear) approximation to h(), we obtain

(1) $$s^* \approx \beta_0 + a\,\beta_a + y\,\beta_y + \gamma\,\beta_\gamma,$$

where $\beta_a = (\partial f/\partial a) + (\partial f/\partial e)*(\partial g/\partial a)$ reflects both the direct effects of ability on achievement and indirect effects operating through the choice of expenditures. The effects of income and tastes operate solely through expenditure choices: $\beta_y = (\partial f/\partial e)*(\partial g/\partial y)$ and $\beta_\gamma = (\partial f/\partial e)*(\partial g/\partial \gamma)$.

It is straightforward to extend the set-up to consider a family that survives for T periods, consuming and earning in each. The utility function is now $u=U(s, c_1, \ldots, c_T; \gamma)$. The test score depends on investment in each period, $s = f(a, e_1, \ldots, e_T)$, where the derivatives $\partial f/\partial e_t$ reflect the marginal productivity of educational investments made in year t. If the family's income in period t is y_t and the interest rate is r, the conventional intertemporal budget constraint (in year-0 dollars) is $\Sigma_t(c_t+e_t)(1+r)^{-t} = \Sigma_t y_t(1+r)^{-t}$. This budget constraint incorporates the implicit assumption that the household can borrow and save freely. If this assumption holds, it is clear that the relevant financial constraint depends only on the discounted value of lifetime income, and not on the family's income in any particular year conditional on this. In other words, the family's consumption and investment decisions are identical to those that would be seen if the family's income were

[3] This can be seen as an indirect utility, where direct utility depends on current and future consumption and where the child's human capital affects future earnings.

[4] Nothing that follows would be changed if the preference parameters entered into the educational production function f() as well, as might be the case if γ reflects variation in (for example) willingness to read to one's children.

constant with the same discounted value, $p = (\Sigma_t y_t(1+r)^{-t}) / (\Sigma_t(1+r)^{-t})$. We refer to p as the family's permanent income.

The family's chosen expenditure path can now be written as $e^* = g(a, p, \gamma)$. Therefore, the direct effect of y_t on the realized test score, controlling for a, p, and γ, will be zero.[5] In other words, we can rewrite (1) as

(2) $s^* \approx \beta_0 + a\,\beta_a + p\,\beta_p + \gamma\,\beta_\gamma$.

C. The interpretation of black-white gaps

Equation (2) is not directly estimable, as a and γ are not readily observed. However, it is useful for interpreting the coefficients of mis-specified regressions with fewer controls. Income is likely to be correlated with ability and attitudes. In an estimate of (2) that omits a and γ, income will explain the black-white test score gap both through its direct effect on the budget constraint and as a proxy for ability and attitudes.[6]

Let b be an indicator for a black student, and let $\delta(X) = E[X \mid b=1] - E[X \mid b=0]$ be the black-white gap in some variable X. By (2), we can write the black-white test score gap as

(3) $\delta(s^*) = \delta(a)\,\beta_a + \delta(p)\,\beta_p + \delta(\gamma)\,\beta_\gamma$.

There are thus three sources of gaps in mean test scores: Differences in incomes ($\delta(p)<0$), differences in ability distributions ($\delta(a)<0$), and differences in preferences ($\delta(\gamma)<0$). The first is importantly different from the latter two: One can easily imagine policy responses (e.g., changes in tax schedules) that would shrink the gap in disposable incomes between black and white families, but it would be difficult or impossible to design policies to reduce differences in ability or preferences between groups.

Now consider a regression that controls for some variable Z as well:

(4) $s^* = \theta_0 + b\,\theta_b + Z\,\theta_Z + u$.

[5] If families face credit constraints or if they are uncertain about future incomes, current and past income may have a larger effect on investment decisions than does future income, potentially creating a role for y in (2).
[6] Mayer (1997) reviews studies that attempt to identify the causal effect of income. See also Dahl and Lochner (2008). This is not our focus here.

By the Frisch-Waugh theorem, the conditional black-white gap, θ_b, is simply the unconditional gap in $\tilde{s}(Z) \equiv s^* - Z\,\theta_Z$; that is, in the component of test scores that is not explained by the control variable Z. This, in turn, can be decomposed into the gaps in the residual portions of ability, permanent income, and preferences, after removing from each the components that are explained by Z:

(5) $\text{plim } \theta_b = \delta(\tilde{s}(Z)) = \delta(\tilde{a}(Z))\beta_a + \delta(\tilde{p}(Z))\beta_p + \delta(\tilde{\gamma}(Z))\beta_\gamma$.[7]

Assuming that Z is positively correlated with a, p, and γ within race and that the β coefficients are all positive, the conditional gap θ_b will be smaller in magnitude than the unconditional gap.

When the control variable Z is permanent income itself, the conditional black-white gap has a simple form. $\tilde{p}(p)$ is identically zero, so

(6) $\theta_b = \delta(\tilde{a}(p))\beta_a + \delta(\tilde{\gamma}(p))\beta_\gamma$.

This is negative only if between-race differences in ability or tastes are larger than would be expected given the within-race association of permanent incomes with ability and tastes and the between-race difference in incomes. Or, more briefly, only if there are differences in ability and tastes between black and white families with the same incomes. Evidence that the black-white test score gap is largely robust to controls for permanent income would therefore suggest that the gap is primarily attributable to black-white differences in ability or preferences/parenting styles rather than to the direct effects of income. By contrast, evidence that the black-white gap was fully eliminated by the inclusion of permanent income controls would be less conclusive, indicating either that the raw gap is attributable primarily to the budget constraint (i.e. β_p is large) or that income is a sufficient proxy for ability and attitudes.

---

[7] It is the within-race relationship between z and {s*, a, p, γ} that must be removed. Algebraically, let B be the matrix formed by concatenating a vector of ones to the race indicator b. Define $M_B = I - B(B'B)^{-1}B'$ – this is the operator that removes race-specific means – and further define $M_{MBZ} = I - Z(Z'M_BZ)^{-1}Z'M_B$. Pre-multiplying a variable X by M_{MBz} effectively removes the component of X that is explained by Z, using the within-race covariance between X and Z. The claims in the text then hold if we define $\tilde{s}(Z) = M_{MBz}s^*$, $\tilde{a}(Z)_z = M_{MBz}a$, and so on.

D. Sensitivity to controls

The black-white test score gap has proven to be surprisingly resilient to controls for income and other family characteristics, for school segregation and integration, and for school spending (Jencks and Phillips 1998). Hedges and Novell (1998), for example, find that differences in parental education and family income explain only about 30% of the black-white gap. Phillips et al. (1998; see also Grissmer and Eiseman 2008) find that broader measures of family environment – including mother's perceived self-efficacy and parenting practices – can explain somewhat more.

The interpretation of the black-white gap conditional on variables that can be seen as the choice of the family – such as the number of books in the home or even characteristics of the school attended – is not straightforward, as these controls may represent both channels by which income effects outcomes and proxies for family preferences. Finding a set of controls for which the conditional gap is zero would provide some insight into the source of the gap, although further inference about the educational production function would be difficult.

In an important recent paper, Fryer and Levitt (2004) control for a relatively short list of covariates, mixing family characteristics and choice variables: A socioeconomic status index (based on parental education and occupational status and on the family's income), gender, the number of children's books in the home, the child's age and birth weight; indicators for teen mothers and for older mothers, and an indicator for receipt of WIC benefits. They find that these variables can fully explain the black-white gap in reading and (to a lesser extent) math scores among entering kindergarteners in the Early Childhood Longitudinal Study (discussed below). However, this result does not persist for long: By the time the same sample of children finished third grade, the raw black-white gap had grown, and only two thirds is explained by these covariates (Fryer and Levitt 2006).

III. Methods

Let s_{it} represent the test score for student i in year t. Let b_i be an indicator for whether the student is black. Let Y_{it} represent the family's income in year t, and let P_i

represent the family's permanent income, defined as the long-run mean of Y_{it}. Let y_{it} and p_i represent the natural logs of Y_{it} and P_i, respectively.

Assuming that test scores are linear in the log of permanent income, we would like to estimate the following regression:

(7) $s_i = \theta_0 + b_i \theta_b + p_i \theta_p + \varepsilon_i$.

Here, $\theta_p > 0$ represents the association between permanent income and test scores, reflecting both causal and non-causal factors. θ_b is the parameter of interest: It represents the black-white gap in scores conditional on permanent income.

Unfortunately, P_i and p_i are not commonly observed in educational data sets. As an alternative, researchers often use annual income as a proxy:

(8) $s_i = \theta_0' + b_i \theta_b' + y_i \theta_y' + \varepsilon_i'$.

In general, $\theta_b' \neq \theta_b$. With a few reasonable assumptions, we can apply the conventional errors-in-variables (EIV) formula to recover the relationship between $(\theta_b\ \theta_p)$ and $(\theta_b'\ \theta_y')$. Let $e_i = y_i - E[y_i \mid p_i]$ represent the transitory component of income. We assume that this is uncorrelated with race: $E[p_i e_i] = E[b_i e_i] = 0$.[8] We also assume that it is uncorrelated with the residual in the test score equation (1), ε.[9] Finally, we assume that the variances of y and e are the same for blacks and whites – that is, that $var(p \mid b)$ and $var(e \mid b)$ are constant across b=0 and b=1. This is not essential, but it keeps the notation simpler. We then define

(9) $R_b \equiv cov(p_i, y_{it} \mid b_i) / var(y_{it} \mid b_i)$.

---

[8] If P_i is defined as the average income for family i within a panel, $(Y_{it} - P_i)$ is necessarily orthogonal to any non-time-varying variable. The assumption need not hold for log incomes, as p_i is not the sample mean of y_{it}. In practice, computing p_i as the sample mean of $\ln(Y_{it})$ rather than as the log of the mean of Y_{it} has little effect on our results. A second issue is that for any particular t (e.g., the year in which a child is ten years old), $Y_{it} - E[Y_{it} \mid P_i]$ might be correlated with ε_{it} if the shape of the lifetime income profile varies with, e.g., ability. We assume that this is not the case, but we evaluate it by controlling for p_i and y_{it} simultaneously. See footnote 19.

[9] This is not innocuous: If the permanent income hypothesis fails, transitory shocks to income might have real, short-term effects on investment decisions and thereby on achievement. Alternatively (see previous footnote), the time path of income might carry information about the family's ability or tastes.

This is the reliability of annual income, viewed as a noisy measure of permanent income, within race. Note that if y is an unbiased measure of p, $R_b \leq 1$, with $R_b=1$ only if $var(e_{it}) = 0$. There are two EIV results. First,

(10) $\theta_y' = R_b \theta_p$.

The income coefficient itself will be attenuated relative to what would be obtained with an error-free permanent income measure. Second,

(11) $\theta_b' = \theta_b + (1-R_b) \delta(p)\theta_p$.

Alternatively, we can re-arrange (11) to write θ_b' as a weighted average of the black-white gap conditional on permanent income, θ_b, and the unconditional gap, $\delta(s) = \theta_b + \delta(p)\theta_p$:

(12) $\theta_b' = \theta_b R_b + (1-R_b) \delta(s)$.

By (11), θ_b' will be smaller (more negative) than θ_b if black families have lower permanent incomes than white families and if annual income is a noisy measure of permanent income. Intuitively, black families will on average have lower permanent incomes than white families with the same annual incomes. As a consequence, controlling only for annual income will lead to overstatement of $|\theta_b|$.

We next develop two methods for recovering θ_b when the data set containing s_{it} lacks a measure of p. We assume that an auxiliary data set is available that contains measures of p, b and y (but not s) for a sample from the same population represented by the primary data set.

A. An IV-based correction

Our first method for recovering the black-white gap controlling for permanent income uses instrumental variables methods. So long as the transitory component of annual income, e, is uncorrelated with the test score error, y is a valid instrument for p in (1). This permits a two-sample two-stage least squares (TS2SLS) estimator. First, we estimate a "first stage" regression of permanent income on current income and race in the auxiliary sample:

(13) $p = \kappa + y \lambda + b \tau + v$.

Note that this is the reverse of the intuitive regression of current on permanent income. In that regression, we would expect a coefficient of one on permanent income and zero on race. In this regression, we expect $0 < \lambda < 1$ and $\tau < 0$. Indeed, it is straightforward to show that plim $\lambda = R_b$ and that plim $\tau = (1\text{-}R_b)\delta(p)$.

The second stage (a.k.a. "reduced form") regression is simply equation (8). Inverting (10) and (11), we can relate the true coefficients to the coefficients from (8): $\theta_p = \theta_y{}' / R_b$ and $\theta_b = \theta_b{}' - (1\text{-}R_b)\, \delta(p)\, \theta_p$. The TS2SLS estimator for $(\theta_p\ \theta_b)$ simply substitutes the first stage coefficients λ and τ into these expressions.[10] Standard errors are obtained by the delta method, using the fact that the first and second stage regressions are estimated on different samples and are therefore statistically independent.

B. Errors-in-variables corrections

An alternative method relies on the standard multivariate errors-in-variables formula (given in, e.g., Greene 2000, p. 377). This relates the coefficients from (8) to the desired coefficients:

$$(14)\quad \begin{pmatrix} \theta_y{}' \\ \theta_b{}' \\ \theta_0{}' \end{pmatrix} = Q^{-1}Q^{*} \begin{pmatrix} \theta_p \\ \theta_b \\ \theta_0 \end{pmatrix}.$$

Here, $Q = E[(y\ \ b\ \ 1)'(y\ \ b\ \ 1)]$ is the design matrix for regression (8), and $Q^* = E[(p\ \ b\ \ 1)'\,(p\ \ b\ \ 1)]$ represents the design matrix of the regression that we would like to run, using permanent income in place of current income. We can recover the desired coefficients by simply inverting (14):

$$(15)\quad \begin{pmatrix} \theta_p \\ \theta_b \\ \theta_0 \end{pmatrix} = \left(Q^{*}\right)^{-1} Q \begin{pmatrix} \theta_y{}' \\ \theta_b{}' \\ \theta_0{}' \end{pmatrix} = Q^{-1}Q^{*}.$$

All that is required are estimates of Q and Q^*. Q is readily estimated from the primary data set. Q^* is nearly identical, except that the upper left entry of Q, $E[y'y]$, is replaced

---

[10] An numerically equivalent approach would be to use the coefficients from (13) to generate a predicted permanent income, $p^{pred} = \kappa + y\,\lambda + b\,\tau$, in the primary data set, then estimate θ_b and θ_p from a regression of s on p^{pred} and b.

by E[p'p]. To recover this, we note that E[p'p] = RE[y'y], where R is the unconditional reliability of current income.

Our EIV estimator thus proceeds in three steps. First, we estimate R using the auxiliary data set, as the R-squared from a regression of y on p. Second, we estimate regression (8) in the primary data set. Finally, using the design matrix from this regression as an estimate of Q, we form Q^* by multiplying the upper left element by our estimate of R and use the result to implement the correction (15).[11]

In practice, the EIV estimator is nearly identical to the TS2SLS estimator.[12] As we see below, the two yield very similar results.

C. Additional control variables

We have developed our two approaches in terms of a model for test scores as functions only of permanent income and race. Researchers often consider specifications with additional control variables – ranging from child age to parental education to the number of books in the home. Our methods can easily be extended to this case; the key required assumption is that the control variables are uncorrelated with current income conditional on permanent income. For non-time-varying controls this is obvious; for others, it may follow from the permanent income hypothesis.

When the control variables are available in both the primary and the auxiliary data sets, the extension of each of our strategies is straightforward. We merely augment each of the regressions above with the additional control variables. The income coefficient is more severely attenuated the richer are the controls, as the covariates will absorb some of the signal but none of the noise, so the corrections have particularly large effects on this coefficient; their effects on the race coefficient are more complex and not easily signed.

---

[11] An alternative algorithm would be to estimate Q^* solely using the auxiliary data set rather than basing it on an estimate of Q from the primary data set. This is the two-sample instrumental variables (TSIV) estimator discussed by Solon and Inoue (2006), who show that TS2SLS is asymptotically more efficient than TSIV. We conjecture that our EIV estimator – which, like TS2SLS but not TSIV, adjusts for differences in Q between the two samples – is also more efficient than TSIV.

[12] The difference between the two can be seen to depend on the source of the estimate of R_b. In both cases, R is estimated from the auxiliary data set. R_b depends on R and on E[(p b)'(p b)]. The TS2SLS estimator uses values from the auxiliary sample, while the the EIV estimator uses the primary sample (and the fact that E[y'b]=E[(p+e)'b)=E[p'b]).

In some of the cases that we consider below, some of the control variables are available only in the primary data set. In this case, only the EIV correction is available, as the reliability of current income conditional on covariates is not consistently estimated via a first stage regression (13) that includes only a subset of the controls.

IV. Data

In this Section, we describe the data used in our analyses. For income data, we use the 1979 National Longitudinal Survey of Youth (NLSY79). Over 12,000 14-22 year olds were first surveyed in 1979 and have been surveyed frequently (annually until 1994 and biennially thereafter) ever since. We use data through 2006, when the youngest respondents were 41 years old.[13] At each survey, respondents are asked detailed questions about their family incomes from various sources. In addition, biological children of female members of the initial sample have been surveyed biennially since 1986, and have been administered standardized tests periodically as they have aged. This sample is known as the "Children of the NLSY," or CNLSY.

The CNLSY testing regime has changed over time, so that (for example) 6-year-olds have scores from different tests depending on the year in which they were born. We focus on three scores are relatively consistently available: The Peabody Individual Achievement Test (PIAT) in math, the PIAT reading recognition and reading comprehension tests (which we average and refer to as a "reading" score), and the Peabody Picture Vocabulary Test-Revised (PPVT-R). We use scores on these three tests from the biennial survey corresponding to the year when the child was 10 or 11, as all CNLSY participants should have been administered these tests at that time, and in most analyses we control for the age at which the exam was taken.[14] Scores on each test are normalized to mean zero and unit variance (based on the CNLSY's 1968 norming sample).

[13] Most respondents from the NLSY's military sample and economically disadvantaged white oversample were dropped from the panel relatively early, and are excluded from our samples. These subsamples represent 0.1% of our main analysis sample.
[14] In some cases, the CNLSY does not seem to have administered the testing protocol perfectly. We allow testing to have taken place anytime after age 9.5 or before age 12.5. We exclude children with no scores in this three-year window.

The NLSY sample is representative of people who were age 14-21 at the end of 1978, so our CNLSY subsample is representative of children born before 1996 to women born between 1958 and 1965. It is not representative of all 10-11 year old children from any particular cohort. Most importantly, children born to older mothers are underrepresented in the CNLSY sample. Accordingly, in most of our analyses of the CNLSY data we control for a quadratic in the mother's age at the child's birth.

We also use a second data set that has been more frequently used for recent studies of the black-white test score gap (e.g., Fryer and Levitt 2004, 2006). The Early Childhood Longitudinal Survey (ECLS) Kindergarten Cohort panel follows a random sample of 21,000 students who were enrolled in kindergarten in the 1998-1999 school year. New data sets have been released about every two years, with data updated to the most recent data collection. We rely on data updated to the spring of students' 5th grade years.[15] Our analysis focuses on students' math scores from the spring of 5th grade (and, in some analyses, from the spring of 3rd grade). We use scaled IRT scores, standardized to have mean zero and unit variance.

Our study focuses on the black-white test score gap. Accordingly, we exclude from our analysis samples of both the CNLSY and ECLS data any respondent who is not either black or non-Hispanic white. Tables 1A and 1B show summary statistics for the two samples. The first column of Table 1A shows statistics for all children of NLSY women who have at least one test score in our age range. The second column represents our sample, and excludes families for which we are unable to construct all relevant income variables.

The NLSY is our primary source of income information. In each survey year, NLSY respondents are asked detailed questions about income from a variety of sources, such as wages and salary, income from self-employment, unemployment insurance, child support, and public benefits. Information about the income of the respondent's spouse (and in later years, his or her resident partner), if any, is also collected. We form family incomes for each year by summing across each of the various components, including income of the spouse if present. To preserve comparability over time, we exclude any

[15] Strictly, the sample is students in 2004 who were in kindergarten in 1999 or first grade in 2000. Most but not all were in 5th grade in 2004.

income from an unmarried partner from the family income calculation. We also censor annual incomes greater than $200,000 (in 1983 dollars), as this censoring point was used in the early waves of the NLSY survey.

We use the family income in the year in which a CNLSY child was tested as his or her current income. To form permanent income, we average the real family income (in 2005 dollars) over the years in which the mother was aged 25 to 39. We also sometimes examine averages over several years prior to the CNLSY test. For these averages, we use only income from even numbered years. That is, we might examine the average of income in the year of the CNLSY test and two years prior, or the average of these two and the income four years prior to the test.

One challenge that we face is that missing values are relatively frequent in the NLSY income questions. In each year, approximately 16% of our sample has missing values in one or more components. If any respondent with a missing component in any survey year were excluded from our permanent income calculation, we would have values for only 29% of CNLSY children. This would be an excessively restrictive rule, as it would exclude (for example) someone with missing food stamp benefit information in a single year even if all other information were complete. Moreover, even observations with complete data from each survey are missing income data for odd-numbered years after 1994, when the NLSY survey became biennial.

To permit consistent measurement of permanent income for as many observations as possible, we developed an extensive imputation algorithm, based loosely on that used by Dahl and Lochner (2008). Where possible, we use information about income of a particular type (e.g., food stamps or child support) from surrounding years to interpolate values for years in which this information is missing. Where there was too much information missing to permit this, we used coarser imputation procedures. Our full imputation algorithm – described in the data appendix – allowed us to form a usable current income for 99% of CNLSY children. We were able to construct a permanent income for 95%, with missing values arising primarily when mothers permanently exited the NLSY sample before age 39. Log current incomes average 10.7 (standard deviation 1.0), while the log permanent incomes average 10.8 (SD 1.0).

The ECLS has much less information about incomes. In surveys administered in the springs of Kindergarten, first, third, and fifth grade, parents of ECLS children were asked the total income of all persons in the household, from all sources. For confidentiality, responses were assigned to 13 bins (except in the kindergarten survey); values were imputed using a hot-deck procedure if the income question was not answered. We assign each bin to its midpoint, using $300,000 for the "$200,000 or more" bin, then convert these values to real 2005 dollars. The current income that we use for analyses of 5th grade test scores is that from the spring of 5th grade, and other covariates come from that wave if possible.[16] We also construct a "short-run" average income by averaging the four annual measures. We set this to missing unless there are at least three non-missing values, two non-imputed.

In some analyses, we also use a socioeconomic status (SES) index provided by the ECLS. This is a composite based on family income, parents' occupation, and parents' education. Because the current family income is used in forming the composite, the SES variable is mechanically correlated with current income, violating the assumptions of our EIV estimator. To eliminate this, we regress SES on current income and use the residual from this regression as a measurement of the non-income components of SES.

V. Permanent and Current Income in the NLSY

In this Section, we present simple analyses of the relationships between race, permanent income, and annual income in the NLSY. Our sample for all analyses is women in the original NLSY sample with children in the CNLSY sample; "current" income is the family income measured in the year that the child was aged 10 or 11, as discussed in the previous Section.

Table 2 presents several simple regressions for annual and long-run income. Column 1 shows that black students' families have log incomes 0.86 below those of white students' families, on average. This gap falls to 0.63 when we control for gender,

---

[16] In some analyses, we use an indicator for WIC receipt. This was asked only the first time a child appeared in the ECLS sample (usually kindergarten, but sometimes later for children in "freshening" samples that were added to the panel in later grades).

child's age, mother's age, year, and birth order (Panel B). Column 2 repeats these specifications for the log of the long-run average annual income. The gap is slightly smaller here.

Column 3 presents a regression of current income on race and long-run average income. When we include our simple demographic controls, the average income coefficient is statistically indistinguishable from one, the black coefficient indistinguishable from zero, and the R^2 just over 0.6. This is consistent with our maintained assumption that the transitory component of current income when children are age 10-11 is uncorrelated with the family's race.

Column 4 reverses this regression, placing long-run average income on the left hand side and current income on the right. Here, the current income coefficient is much smaller, just above 0.5. The black coefficient is negative, around -0.3, and highly significant. This demonstrates the central fact that underlies our analysis: Even when current incomes are controlled, the black-white gap in permanent income remains substantial. Indeed, the residual gap is just a bit less than half as large as the unconditional black-white permanent income gap, and it implies that test score regressions that include current income as an explanatory variable will dramatically understate – by nearly half – the explanatory power of family income for the black-white test score gap.

Columns 5 and 6 repeat the specifications from Columns 3 and 4, this time excluding the race control. Unsurprisingly, the income coefficients are largely unchanged. The R^2 from the model in Panel A, Column 5 is 0.615. We use this below as our estimate of the reliability of current income as a measure of permanent income, R.

VI. Results

In this Section, we present our main results for the black-white test score gap, first using the CNLSY data and then the ECLS sample.

A. Evidence from the CNLSY

Table 3 presents regressions for student scores on the PIAT math exam, given to children of the original NLSY respondents when the children are 10 or 11 years old. Column 1 presents an analysis using the maximal possible sample and with no controls.

The raw black-white gap is 0.77 standard deviations. Column 2 (and the remainder of the table) restricts the sample to families for whom we observe enough information to compute a permanent income as well as annual income in the year of the test, two years prior, and four years prior. The gap in this subsample is nearly identical, 0.76.

Column 3 adds the vector of demographic controls used in Table 2: Child gender, the child's age at the time of the exam and its square, the mother's age at the child's birth and its square, the child's parity (entered as dummy variables), and calendar year indicators.[17] These controls bring the gap down to 0.56; a substantial portion of the raw black-white gap is attributable to differences in the distribution of mother's age at the child's birth between blacks and whites.

Column 4 adds a control for contemporaneous log family income. This has coefficient 0.21, indicating that a 10% increase in family income is associated with an increase in student test scores of about 0.02 standard deviations. The black coefficient shrinks to -0.43, about one quarter smaller than in Column 3.

Columns 5-7 present specifications that use alternative income measures: The average of current income and that two years prior (Column 5); the average of current, two years prior, and four years prior (Column 6); and our long-run average (Column 7). As expected, when we use more information to construct our income measures, the income coefficient gets larger and the black coefficient shrinks toward zero. In Column 7, the black coefficient has fallen to -0.36, down 15% from that in Column 4.

Columns 8 and 9 present estimates from our two two-sample techniques for recovering the black-white gap conditional on permanent income. In Column 8, we report estimates from an instrumental variables specification, using the log of current income as an instrument for the log of the long-run average income. (The reduced form for this specification is the model in Column 4.) In Column 9, we apply the parametric errors-in-variables correction discussed in Section III to the estimates from Column 4, assuming that annual income has reliability R=0.615.[18] Estimates in Columns 8 and 9

---

[17] We have estimated all specifications without the controls and including observations for which some of the income lags are missing where possible, with similar results.
[18] The standard errors here treat the reliability of annual income as known with certainty.

are quite similar. The income coefficients are notably larger and the black coefficients are notably smaller than in Column 7.[19]

Taking the estimates in Table 3 together, it is clear that simply including annual income in a regression severely under-controls for differences in permanent income between black and white families. The black coefficient in Column 9 is only 70% as large as that in Column 4. Stated somewhat differently, the inclusion of an annual income control explains only 51% as much of the raw black-white gap (as in Column 3) as is explained by permanent income.

Table 4 presents estimates for all three of the test scores available in the NLSY. The raw black-white gap is much larger on the PPVT than on the PIAT math, and is somewhat smaller on the PIAT reading. But the general pattern as we compare different income controls is, not surprisingly, very similar: Controlling for current income gets us only about half way to the black-white gap conditional on permanent income.

B. Evidence from the ECLS

Table 5 presents estimates for students' math scores in the ECLS-K, as measured in the Spring of 5th grade, when most students were 10 or 11 years old. Column 1 presents a regression that includes only a single independent variable, the student's race. The raw black-white gap in the ECLS data is 0.85 standard deviations. This is reduced slightly, to 0.78, when we restrict the sample to observations for which we have non-missing data on family income and the mother's age.

Column 3 adds controls for the child's gender and age (entered as a quadratic). These have essentially no effect on the black-white gap. Column 4 adds additional quadratic controls for the mother's age at the child's birth. These are necessary for our

---

[19] One potential explanation is that measurement error in our permanent income variable biases the coefficient in Column 7. This should bias the 2SLS and EIV estimates in Columns 8-9 as well. But if income in other years is measured with greater error than is current income – as is plausible given that current income is not imputed in this sample but income in other years is in many cases – then the bias in Column 7 would be larger than that in Columns 8 and 9. An alternative explanation for the results is that current income (or income at short lags) has a direct effect on test scores. Although it does not distinguish among the two potential explanations, we have also estimated an OLS specification with both permanent and current income. The income coefficients are 0.240 (0.037) and 0.079 (0.026), respectively, and the black coefficient is -0.360 (0.038).

two-sample analyses, as the CNLSY sample is only representative conditional on the mother's age. The maternal age control explains a notable portion of the gap, shrinking it to 0.62.

Column 5 adds a control for the family's income in the year that the test was taken. This reduces the black-white gap by 0.24, to 0.38. Column 6 replaces the current income control with the average of family income across all four ECLS survey waves. The income coefficient is notably larger here, and the black-white gap shrinks to 0.34.

In Column 7, we present two sample two-stage least squares estimates that use the NLSY data to estimate the first-stage relationship between permanent income and the instrument, current income, and use the ECLS data to identify the reduced-form relationship between current income and test scores (as in Column 5). The identifying assumption here is that the transitory component of income is uncorrelated with both race and achievement. Estimates yield much larger income coefficients than are seen in earlier columns, and the black-white gap is shrunk to 0.20, less than a third of the raw gap and half of the gap controlling for current income.

In Column 8, we use our parametric errors-in-variables correction based on the estimates in Column 5. As before, we use our NLSY-based estimate that the reliability of annual income (viewed as a measure of permanent income) is 0.615.[20] This yields quite similar results to those obtained with the TS2SLS correction.

These estimates almost certainly undercorrect for the role of true permanent income. In both cases, we assume that the current income measure in the ECLS is equivalent to that in the NLSY, when in practice the former is much inferior and likely less reliable. If so, the income coefficients in Columns 7 and 8 remain somewhat attenuated, and the black coefficients somewhat negatively biased.

[20] Haider and Solon (2006) and Mazumder (2003) emphasize that the reliability is likely to vary with age composition of the sample. Unfortunately, differences in the sampling schemes of the two surveys mean that the age distribution of mothers differs: 9.2% of mothers in the ECLS and 2.2% of mothers in the CNLSY sample gave birth after age 35. When we reweight the ECLS sample to match the mother's age distribution in the CNLSY sample, race and income coefficients do not change substantially. While reliability may also differ across cohorts, we have no way to address this. In our defense, the ECLS mothers come from generally similar birth cohorts as those in the NLSY.

C. Additional controls

It is common when analyzing the conditional black-white test score gap to control for other factors in addition to family income. For example, Phillips et al. (1998) explore controls like parental occupational status, parental wealth, neighborhood average income, and variables capturing the quality of the school and home environment. Some of the common controls (e.g., parental education and the presence of a father) may be better proxies for the family's permanent income than is current family income, suggesting that estimates that include these controls should go part of the way toward correcting the biases that are the focus of this study. We can use our methods to investigate whether simple controls can adequately address the problem.

Table 6 presents estimates that control for two widely-available and commonly controlled variables that are plausibly good measures of permanent family income, maternal education and for the presence of a father.[21] The first panel presents estimates from the CNLSY, while the second presents estimates from the ECLS. Both of the new variables are available in each sample. Column 1 presents estimates without income controls, Column 2 adds the current income, Column 3 uses an average income over a longer period instead, and Columns 4 and 5 present estimates using our 2SLS (TS2SLS in Panel B) and EIV corrections, respectively. Somewhat surprisingly, the black-white gap is a bit larger with the additional controls than in analogous specifications without them (in Tables 3 and 5). But the inclusion of these controls has only a moderate effect on our main results; even when maternal education and family structure are controlled, a model with current family income overstates the conditional black-white gap by 29 (NLSY) to 90 (ECLS) percent relative to what would be obtained were permanent income controlled via our EIV estimator.

As a final exercise, we explore the implications of our analysis for Fryer and Levitt's (2006; hereafter FL) investigation of the black-white test score gap among 3rd graders in the ECLS. In an earlier paper (Fryer and Levitt 2004), these authors showed that a relatively small set of covariates could fully explain the black-white test score gap

[21] We use only a contemporaneous family structure variable here. We have also explored specifications that control for the fraction of the child's life in which the father was present, with similar results.

in kindergarten; in the 2006 paper they find that this is no longer true by the time the same students finish 3rd grade. Columns 1 and 2 of Table 7 report their estimates of the raw gap and the gap conditional on a list of nine covariates, ranging from the child's age and birth weight to measures of mother's age to the number of children's books in the home.[22] The raw gap is 0.88 standard deviations, and the inclusion of the FL controls reduces this to 0.38.

We have not been able to fully replicate FL's sample.[23] We nevertheless obtain very similar results, as seen in Columns 3 and 4.[24] Columns 5 and 6 repeat the estimates on the subsample of students for whom we have non-missing, non-imputed family income. The black-white gap, both unconditional and conditional on the FL covariates, is notably smaller in this subsample, but the conditional gap remains large and significant.

Column 7 adds the log of current family income to the specification. For ease of interpretation, we make one adjustment to the FL covariates: We replace the socioeconomic status index, which depends mechanically on current income, with a residualized variable that measures the portion of the SES index that is orthogonal to log current income. Both log income and the residual component of SES are strong predictors of children's test scores, and the income coefficient is quite similar to that seen in Column 5 of Table 6.

Because not all of the FL variables are available in the NLSY, we cannot apply our TS2SLS correction to the specification in Column 7 of Table 8. However, our EIV correction can be used here. As before, the key identifying assumption is that the transitory component of current income is uncorrelated with any of the control variables.

---

[22] FL's analysis differs from ours in including racial groups other than blacks and non-Hispanic whites in the analysis. We do not report their coefficients on indicators for the other groups in Table 7. In the estimates in Columns 3-8, we exclude members of those groups from our sample.

[23] The most likely explanation for our failure to replicate their results is that we use the 5th grade wave of the ECLS data where they (presumably) used the 3rd grade wave. Students who attrited from the survey after 3rd grade are missing from our sample. Also, we take all control variables from the 3rd grade survey where possible; we believe that FL use the kindergarten survey as the source of most of their covariates. Finally, we use the ECLS's 3rd grade cross-sectional weights, where FL use longitudinal weights.

[24] Our heteroskedasticity-robust standard errors are somewhat larger than those of FL. FL appear to have used classical SEs; when we do the same, ours closely resemble theirs.

This is not obvious here: Current income may be correlated with the number of books in the home, even conditional on the family's permanent income. (Note, however, that there would be no correlation if the household behaved according to the permanent income hypothesis and faced no credit constraints.) Nevertheless, it seems likely to be a reasonably accurate approximation.

Applying the EIV correction, in Column 8, we see that the income coefficient is increased by a factor of nearly three relative to Column 7. This has dramatic effects on the coefficients for the other covariates: The children's books and maternal age coefficients cease to be significant; the variable measuring WIC receipt switches signs, indicating that conditional on permanent income WIC recipients earn much higher test scores than non-recipients; and the coefficient on the residual component of the SES index increases by a third. Finally, the black coefficient is dramatically shrunken, to 0.18 standard deviations, just over half of the estimate from an otherwise-identical specification without income controls (in Column 6).[25] Evidently, permanent income and the other SES components can explain nearly all of the black-white test score gap, without even resorting to proxies for parenting choices like the number of books in the home.

VII. Discussion

Researchers commonly examine the black-white test score gap while controlling for covariates like family income. The interpretation of these regressions is not straightforward. The income coefficients cannot be interpreted as the causal effect of income on student achievement, as they may capture other family characteristics – including genetic endowment, parenting styles, preferences, etc. – that are themselves important determinants of student achievement but are difficult to measure and control directly. Thus, the black coefficient from these regressions cannot be interpreted as the amount by which black students would underperform relative to whites after a policy intervention that equalized family incomes but not the other characteristics that load onto income.

[25] This is somewhat sensitive to the choice of weights. When we use longitudinal weights, the black coefficient is -0.09 and is statistically insignificant.

Nevertheless, regressions of this form can be interesting. As Phillips et al. (1998, p. 104) write, these regressions "can give us some sense of which characteristics are likely to be important." In particular, it is natural to interpret the extent to which the black coefficient is shrunk by the inclusion of an income control as an upper bound to the effect of differences in black and white adults' labor market outcomes on their children's achievement.

Previous researchers have found that family income and other variables measuring a family's external circumstances do a relatively poor job of explaining the black-white test score gap. However, the literature has ignored a key factor: There is little theoretical justification for believing that current income, rather than permanent income, is an important determinant of student achievement, and empirically current income turns out to be a very poor proxy for longer-run measures of families' financial circumstances.

We develop two methods for identifying the black-white test score gap conditional on permanent income. These methods can be used even when the data set containing student test scores does not itself permit accurate measurement of a family's permanent income. They would also be useful for examinations of racial gaps in other outcomes such as educational attainment, asset accumulation (Hurst, Luoh, and Stafford 1998; Mayer 1997), and consumption patterns (Charles, Hurst, and Roussanov, forthcoming).

We apply our estimators to samples from the CNLSY and the ECLS data sets. In both samples, the black-white test score gap conditional on permanent income is much smaller than the gap conditional on current income, and a comparison to the unconditional gap indicates that family financial circumstances can explain 40 to 75% of the raw black-white test score gap at age 10 or 11. Moreover, we find that the addition of a control for permanent income to the already-rich covariates considered by Fryer and Levitt (2006) halves the unexplained gap. Other variables – like maternal education, the presence of a father, or occupation-based socioeconomic status indices – do not do nearly as good of a job of capturing the family circumstances that are related with student achievement and that differ between races.

It is important to be cautious in interpreting our results, as we cannot tell whether the large apparent effects of permanent income reflect the true causal effects of family resources or other, less manipulable factors that are correlated with economic outcomes. Nevertheless, our results at least hold out some hope for the possibility that improvements in black families' economic circumstances, absent any other changes, could lead to substantial closing of the black-white test score gap.

References

Blau, David M. (1999). "The Effect of Income on Child Development," *Review of Economics and Statistics* 81(2), May, 261-276.

Blau, Francine D. and Adam J. Grossberg (1992). "Maternal Labor Supply and Children's Cognitive Development," *Review of Economics and Statistics* 74 (3), August, 474-481.

Brooks-Gunn, Jeanne, Pamela Klebanov, and Greg J. Duncan (1996). "Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics," *Child Development* 28(6), 1048-1055.

Campbell, Mary E., Robert Haveman, Tina Wildhagen, and Barbara L. Wolfe (2008). "Income Inequality and Racial Gaps in Test Scores." In Magnuson, Katherine and Jane Waldfogel, eds., Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap. Russell Sage Foundation: New York, 110-136.

Charles, Kerwin, Eric Hurst, and Nick Roussanov (forthcoming). "Conspicuous Consumption and Race." Forthcoming, *Quarterly Journal of Economics*.

Dahl, Gordon and Lance Lochner (2008). "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." Unpublished manuscript, Nov. 21.

Friedman, Milton (1957). A Theory of the Consumption Function. National Bureau of Economic Research: New York.

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* 86(2), May, 447-464.

Fryer, Roland G., Jr. and Steven D. Levitt (2006). "The Black-White Test Score Gap Through Third Grade," *American Law and Economics Review* 8(2), Summer, 249-281.

Greene, William H. (2000). Econometric Analysis 4th ed. Prentice Hall: Upper Saddle River, New Jersey.

Grissmer, David and Elizabeth Eiseman (2008). "Can Gaps in the Quality of Early Environments and Noncognitive Skills Help Explain Persisting Black-White Achievement Gaps?" In Magnuson, Katherine and Jane Waldfogel, eds., Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap. Russell Sage Foundation: New York, 139-180.

Haider, Steven and Gary Solon (2006). "Life-Cycle Variation in the Association Between Current and Lifetime Earnings," *American Economic Review* 96, September, 1308-1320.

Hedges, Larry V. and Amy Nowell (1998). "Black-White Test Score Convergence since 1965." In Jencks, Christopher and Meredith Phillips, ed., The Black-White Test Score Gap, Brookings Institution Press: Washington DC, 149-181.

Herrnstein, Richard J. and Charles Murray (1996). The Bell Curve: Intelligence and Class Structure in American Life. Free Press: New York.

Hurst, Eric, Ming Ching Luoh, and Frank Stafford (1998). "Wealth Dynamics of American Families: 1984-1994," *Brookings Papers on Economic Activity* 1.

Jencks, Christopher and Meredith Phillips (1998). "The Black-White Test Score Gap: An Introduction." In Jencks, Christopher and Meredith Phillips, ed., The Black-White Test Score Gap, Brookings Institution Press: Washington DC, 1-55.

Kozol, Jonathan (1991). Savage Inequalities: Children in America's Schools. Crown Publishers: New York.

Krueger, Alan, Jesse Rothstein, and Sarah Turner (2006). "Race, Income, andCollege in 25 Years: Evaluating Justice O'Connor's Conjecture," *American Law and Economics Review* 8(2), Summer, 282-311.

Lareau, Annette (2003). Unequal Childhoods: Class, Race, and Family Life. University of California Press: Berkeley CA.

Magnuson, Katherine and Jane Waldfogel (2008). "Introduction." In Magnuson, Katherine and Jane Waldfogel, eds., Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap. Russell Sage Foundation: New York, 1-30.

Mayer, Susan E. (1997). What Money Can't Buy: Family Income and Children's Life Chances. Harvard University Press: Cambridge, MA.

Mazumder, Bhashkar (2001). "The Mismeasurement of Permanent Earnings: New Evidence from Social Security Earnings Data." Federal Reserve Bank of Chicago Workign Paper 2001-24.

Mazumder, Bhashkar (2003). "Revised Estimates of Intergenerational Income Mobility in the United States." Federal Reserve Bank of Chicago Workign Paper 2003-16.

Modigliani, F. and R. Brumberg (1954). "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data." In Kurihara, Kenneth K., ed., Post-Keynesian Economics, Rutgers University Press: New Brunswick NJ.

Moynihan, Daniel Patrick (1965). The Negro Family: The Case for National Action. United States Department of Labor, Office of Policy Planning and Research: Washington DC.

Neal, Derek (2006). "Why Has Black-White Skill Convergence Stopped?" In Hanushek, Eric and Finis Welch, Handbook of Economics of Education. Elsevier: Amsterdam.

Phillips, Meredith, Jeanne Brooks-Gunn, Greg J. Duncan, Pamela Klebanov, and Jonathan Crane (1998). "Family Background, Parenting Practices, and the Black-White Test Score Gap." In Jencks, Christopher and Meredith Phillips, ed., The Black-White Test Score Gap, Brookings Institution Press: Washington DC, 103-148.

Solon, Gary (1992). "Intergenerational Income Mobility in the United States," *American Economic Review* 82, June, 393-408.

Solon, Gary and Atsushi Inoue (2006). "Two-Sample Instrumental Variables Estimators." Working paper, December.

Data Appendix

In this appendix, we describe the imputation procedure that we use to fill in missing values in the NLSY income variables. Our procedure is based loosely on that used by Dahl and Lochner (2005), who generously provided us with their programs.

We divide the family's income into 19 components that are reasonably consistently measured in the NLSY. The most important are own wage and salary, spouse's wage and salary, military income for the respondent and for the spouse, self employment income for the respondent and for the spouse, and income "from all other sources," but there are also components reflecting various categories of government transfers (unemployment insurance, welfare, food stamps, SSI, etc.), as well as alimony, child support, and gifts.

We attempt to impute missing values for each of these separately. Wage and salary income, which accounts for 77% of total income in our sample, is quite variable across years for many individuals. Much of this variation appears to come from changes in employment status, so we treat employment status – measured as annual weeks worked – and annual full-year-equivalent earnings as distinct sources of variation, imputing the two separately and then multiplying them together. Similarly, we impute marital status and spouse's age separately, and impute values for the spouse's income only if the respondent appears to have been married in the relevant year.

We use the following strategy to impute full-year-equivalent wage and salary income, military income, self-employment income, "other" income, and the corresponding components for the spouse. If there are five or more non-missing values for a specified component for an individual, we estimate an individual-level regression using all non-missing values, with the respondent's (or her spouse's) age and its square as explanatory variables. We then impute missing values using the fitted values from this regression. If fewer than five non-missing values are available, or if the fitted value from the individual-level regression is negative, we instead impute with fitted values from a global regression that uses all individuals in the sample and includes individual fixed effects along with a single quadratic age control.

Information on employment status is available weekly for all years, even if a survey was not conducted. We linearly interpolate to fill in missing values of the fraction of the year the respondent (or spouse) was employed, using data from the year before and the year after the missing observation. We do not extrapolate employment status or interpolate across gaps greater than three years, so wage and salary income cannot be imputed in these years.

For the other income components, we use a simpler procedure: We simply impute the person-specific mean. We do not impute values if there are fewer than three non-missing values for the component.

If we are able to produce actual or imputed values for wages and salary, military income, and self-employment income, we form total family income as the sum of all available income components, using imputed values when actual values are unavailable and assigning zero to components that cannot be imputed. If we are unable to impute any of these three primary income categories, however, we revert to interpolating family income itself using fitted values from a person-specific regression of total family income on age and its square. We do not extrapolate family income to years outside of the range for which we have actual values. Moreover, if we have to interpolate family income for more than two of the years used in the permanent income calculation, we treat permanent income as missing.

We censor annual family income at $200,000 in 1983 dollars, corresponding to $390,000 in dollars of 2005. We then convert family income to 2005 dollars, and average it over all years when the respondent is aged 25-39 to form our permanent income measure.

Table 1A. NLSY summary statistics

	Population (1)	Sample (2)	Black (3)	White (4)
PIAT math score	0.31	0.33	-0.27	0.48
	[0.98]	[0.98]	[0.94]	[0.92]
PIAT reading score	0.27	0.28	-0.22	0.41
	[0.88]	[0.88]	[0.88]	[0.83]
PPVT score	-0.11	-0.08	-1.09	0.18
	[1.27]	[1.27]	[1.22]	[1.14]
Math or Reading fraction missing	0.02	0.02	0.04	0.02
PPVT fraction missing	0.08	0.08	0.08	0.08
Female	0.49	0.49	0.50	0.49
Child age (months)	130.4	130.5	130.5	130.5
	[7.7]	[7.6]	[7.6]	[7.6]
ln(current family income) at test date	10.70	10.72	10.03	10.90
	[0.97]	[0.97]	[0.95]	[0.89]
Fraction imputed (excluding missing)	0.14	0.14	0.19	0.13
ln(mean family income over ages 25-39)	10.77	10.77	10.17	10.93
	[0.71]	[0.7]	[0.7]	[0.61]
Fraction of years imputed (excluding missing)	0.34	0.33	0.38	0.32
Fraction of years imputed (survey years only)	0.14	0.14	0.19	0.12
Mother's education (years)	13.08	13.18	12.32	13.40
	[2.58]	[2.43]	[2.34]	[2.41]
Father present?	0.76	0.77	0.46	0.84
Mother's age at child's birth	25.85	25.97	23.48	26.62
	[5.22]	[5.19]	[5.2]	[4.99]
Observations	5440	4966	2061	2905

Notes: Table displays means; standard deviations are in brackets. Calculation use CNLSY panel weights. "Population" includes all black and white CNLSY children with at least one test score at age 10 or 11. The "sample" consists of those observations which also have current and permanent income measures.

Table 1B. ECLS summary statistics

	Population (1)	Sample (2)	Black (3)	White (3)
Reading score	0.09	0.19	-0.39	0.31
	[0.98]	[0.96]	[0.95]	[0.92]
Math score	0.07	0.18	-0.46	0.32
	[0.99]	[0.96]	[0.92]	[0.91]
Female	0.48	0.48	0.48	0.48
	[0.5]	[0.5]	[0.5]	[0.5]
Age at test (months)	133.5	133.4	132.8	133.5
	[4.4]	[4.2]	[4.1]	[4.3]
ln(income)	10.80	10.90	10.15	11.05
	[0.97]	[0.95]	[1.03]	[0.85]
ln(mean income)	10.89	10.93	10.25	11.08
	[0.81]	[0.8]	[0.78]	[0.72]
Mother's education (years)	13.66	13.82	12.95	14.00
	[2.28]	[2.27]	[1.77]	[2.33]
Mother's age at child's birth	27.61	27.91	24.66	28.59
	[5.93]	[5.76]	[5.93]	[5.49]
Father present?	0.71	0.76	0.42	0.83
Observations	7731	6138	776	5362

Notes: Table displays means; standard deviations are in brackets. Statistics are weighted using the ECLS fifth grade cross-sectional weights. "Population" includes all black and white ECLS-K respondents in the fifth grade survey with test scores. "Sample" is restricted to respondents for which income measures and mother's age and education are available. Test scores, age, and income are measured in the fifth grade survey. Mean test scores are nonzero because the norming sample includes other races.

Table 2. Racial gaps in income and permanent income

	Dependent variable Ln(current income) (1)	Ln(permanent income) (2)	Ln(current income) (3)	Ln(permanent income) (4)	Ln(current income) (5)	Ln(permanent income) (6)
Panel A. No controls						
Black	-0.858	-0.756	-0.053	-0.306		
	(0.039)	(0.031)	(0.028)	(0.024)		
Ln(permanent income)			1.065		1.078	
			(0.024)		(0.020)	
Ln(current income)				0.524		0.571
				(0.015)		(0.013)
R2	0.129	0.190	0.616	0.642	0.615	0.615
Panel B. With controls						
Black	-0.632	-0.594	-0.020	-0.274		
	(0.040)	(0.033)	(0.027)	(0.022)		
Ln(permanent income)			1.031		1.035	
			(0.026)		(0.023)	
Ln(current income)				0.505		0.537
				(0.016)		(0.015)
R2	0.214	0.272	0.623	0.651	0.623	0.630

Notes: Sample is children in the CNLSY sample for whom tests were administered at age 10 or 11 and for whom both contemporaneous and permanent family income could be constructed. N=5,086. Specifications in lower panel include controls for gender, a quadratic in the child's age at test date, a quadratic in the mother's age at the child's birth, child birth order dummies, and calendar year dummies. Regressions are weighted using a longitudinal weight for the child. SEs, clustered on the mother, in parentheses.

Table 3. Sensitivity of black-white gap on math PIAT scores to alternative income controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	2SLS	EIV
Black	-0.767	-0.758	-0.557	-0.426	-0.410	-0.404	-0.362	-0.318	-0.300
	(0.033)	(0.034)	(0.035)	(0.037)	(0.037)	(0.037)	(0.038)	(0.041)	(0.043)
Ln(current income)				0.205					
				(0.019)					
Ln(avg. of current & 2-yr. lagged income)					0.235				
					(0.022)				
Ln(avg. of current, 2-yr., & 4-yr. lagged income)						0.249			
						(0.023)			
Ln(long run avg. income)							0.327		
							(0.027)		
Ln(permanent income)								0.400	0.402
								(0.037)	(0.038)
Controls?	n	n	y	y	y	y	y	y	y
R2	0.10	0.10	0.18	0.21	0.21	0.21	0.22		

Notes: N = 5,410 in Column 1, 4,942 in 2-9. Sample in columns 2-9 excludes children with missing family income variables. SEs, clustered on the mother, in parentheses. See notes to Table 2 for control variables and weights.

Table 4. Sensitivity of black-white gaps in NLSY to alternative income controls

Method	OLS	OLS	OLS	OLS	2SLS	EIV
Income measure	None	None	None	Contemp.	Perm.	Perm.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: PIAT Math						
Black	-0.767	-0.758	-0.557	-0.426	-0.318	-0.300
	(0.033)	(0.034)	(0.035)	(0.037)	(0.041)	(0.043)
Ln(income) coefficient				0.205	0.400	0.402
				(0.019)	(0.037)	(0.038)
Controls	n	n	y	y	y	y
R2	0.10	0.10	0.18	0.21		
Panel B: PIAT Reading Composite (0.5*reading comprehension +)						
Black	-0.621	-0.625	-0.464	-0.338	-0.232	-0.217
	(0.032)	(0.033)	(0.035)	(0.037)	(0.041)	(0.043)
Ln(income) coefficient				0.197	0.386	0.385
				(0.019)	(0.037)	(0.037)
Controls	n	n	y	y	y	y
R2	0.08	0.08	0.15	0.19		
Panel C: PPVT						
Black	-1.278	-1.278	-1.059	-0.909	-0.787	-0.765
	(0.046)	(0.048)	(0.047)	(0.049)	(0.054)	(0.057)
Ln(income) coefficient				0.235	0.453	0.460
				(0.027)	(0.051)	(0.053)
Controls	n	n	y	y	y	y
R2	0.17	0.17	0.22	0.24		

Notes: See notes to Tables 2 and 3. N = 5,410, 5,325, and 5,048 in Column 1 of Panels A, B, and C, respectively. Columns 2-6 exclude families with missing current or permanent income; N = 4,942, 4,864, and 4,623. SEs, clustered on the mother, in parentheses.

Table 5. Sensitivity of black-white gap on ECLS 5th grade math scores to alternative income controls

Sample	Full	Analysis						
Method	OLS	OLS	OLS	OLS	OLS	OLS	TS2SLS	EIV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.846	-0.775	-0.779	-0.623	-0.383	-0.343	-0.195	-0.151
	(0.025)	(0.066)	(0.065)	(0.071)	(0.067)	(0.070)	(0.077)	(0.077)
Ln(current income)					0.334			
					(0.030)			
Ln(short-run avg. income)						0.434		
						(0.035)		
Ln(permanent income)							0.649	0.656
							(0.062)	(0.059)
Gender, age	n	n	y	y	y	y	y	y
Mom's age at birth	n	n	n	y	y	y	y	y
R2	0.13	0.09	0.10	0.14	0.23	0.24		

Notes: Analysis sample is children in the ECLS sample for whom tests were administered in the spring 5th grade survey. N = 7,742. In Columns 2-8, children for whom family income, average income, or any controls are unavailable are excluded; N = 6,143. Specifications in Columns 3-8 include controls for gender and a quadratic in the child's age at test date. Specifications in Columns 4-8 also include a quadratic in the mother's age at the child's birth. Regressions are weighted by the ECLS 5th grade cross-sectional weight; robust SEs are in parentheses.

Table 6. With controls for maternal education and the presence of a father figure

Method	OLS	OLS	OLS	2SLS	EIV
Income measure	None	Current	Long-run avg.	Permanent	Permanent
	(1)	(2)	(3)	(4)	(5)
Panel A: NLSY (PIAT Math)					
Black	-0.508	-0.457	-0.407	-0.371	-0.354
	(0.035)	(0.037)	(0.038)	(0.043)	(0.044)
Ln(income) coefficient		0.144	0.238	0.322	0.439
		(0.023)	(0.032)	(0.051)	(0.069)
Mother's education	0.095	0.075	0.067	0.058	0.033
	(0.008)	(0.008)	(0.008)	(0.010)	(0.012)
Father present	0.123	0.001	0.005	-0.038	-0.250
	(0.036)	(0.040)	(0.039)	(0.043)	(0.068)
N	4,926	4,926	4,926	4,926	4,926
Panel B: ECLS (Math)					
Black	-0.469	-0.396	-0.362	-0.274	-0.208
	(0.070)	(0.068)	(0.071)	(0.074)	(0.077)
Ln(income) coefficient		0.225	0.295	0.454	0.803
		(0.032)	(0.036)	(0.067)	(0.115)
Mother's education	0.134	0.101	0.096	0.101	0.016
	(0.009)	(0.009)	(0.009)	(0.009)	(0.017)
Father present	0.231	0.042	0.062	-0.006	-0.445
	(0.061)	(0.064)	(0.061)	(0.068)	(0.111)
N	6,143	6,143	6,143	6,143	6,143

Notes: Estimates in Panels A and B include the same controls as in Table 3 and Table 5, respectively. The additional controls are mother's education years and a dummy for whether a father (or stepfather) is present in the household. In Column 4, Panel B shows our TS2SLS estimate. Robust standard errors (clustered on the mother in Panel A) are in parentheses. Regressions in each panel use sampling weights.

Table 7. Analysis of ECLS 3rd grade math scores

	Fryer and Levitt (2006)		Our analysis					
			Full sample		Income subsample			
			OLS	OLS	OLS	OLS	OLS	EIV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.882	-0.382	-0.824	-0.396	-0.776	-0.344	-0.324	-0.180
	(0.031)	(0.033)	(0.031)	(0.033)	(0.038)	(0.039)	(0.039)	(0.042)
Ln(current income)							0.298	
							(0.017)	
Ln(permanent income)								0.801
								(0.047)
Socioeconomic status measure		0.288		0.371		0.382		
		(0.015)		(0.018)		(0.020)		
SES measure (residualized)							0.326	0.437
							(0.026)	(0.028)
Age (months)		0.019		0.020		0.022	0.022	0.025
		(0.002)		(0.003)		(0.003)	(0.003)	(0.003)
Birth weight (oz)		0.003		0.004		0.004	0.004	0.004
		(0.000)		(0.001)		(0.001)	(0.001)	(0.001)
Female		-0.175		-0.139		-0.164	-0.161	-0.146
		(0.018)		(0.022)		(0.024)	(0.024)	(0.024)
# of children's books		0.006		0.002		0.003	0.003	-0.001
		(0.001)		(0.001)		(0.001)	(0.001)	(0.001)
# of books squared (*1000)		-0.020		-0.006		-0.008	-0.007	0.005
		(0.003)		(0.004)		(0.004)	(0.004)	(0.004)
Mother over 30 at 1st birth		0.083		0.097		0.088	0.091	-0.029
		(0.024)		(0.031)		(0.033)	(0.033)	(0.035)
Mother a teenager at 1st birth		-0.132		-0.163		-0.155	-0.153	-0.037
		(0.025)		(0.033)		(0.037)	(0.037)	(0.038)
Mother receives WIC benefits		-0.208		-0.180		-0.162	-0.141	0.220
		(0.024)		(0.030)		(0.033)	(0.034)	(0.046)
N	11,201	11,201	9,934	9,934	8,076	8,076	8,076	8,076
N (black or white non-hisp.)	7,908	7,908	9,934	9,934	8,076	8,076	8,076	8,076
R2	0.12	0.26	0.12	0.28	0.09	0.27	0.27	

Notes: Columns 3 and 4 use our full sample, which excludes children who attrited before the 5th grade interview. Columns 5-8 also require that income is present. Columns 7 and 8 replace the SES control with a control for the component of the SES measure that is orthogonal to income. Columns 3-8 use the ECLS 3rd grade cross-sectional weights. Robust SEs in parentheses.