Permanent Income and the Black-White Test Score Gap

Jesse Rothstein[*]
University of California, Berkeley and NBER

Nathan Wozny
U.S. Air Force Academy

June 2012
First draft: June 2009

Abstract

Analysts often examine the black-white test score gap conditional on family income. Typically only a current income measure is available. We investigate the importance of the income measure to the conditional black-white gap, and find that the gap conditional on observed annual income dramatically overstates the gap conditional on richer measures such as permanent income. We describe a method for identifying the latter using an auxiliary data set to estimate the relationship between current and permanent 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 two commonly used data sets. When we add permanent income to the controls used by Fryer and Levitt (2006), the unexplained gap in 3rd grade shrinks below 0.15 SDs, less than half of what is found with their controls.

I. Introduction

The black-white test score gap has been extensively documented. Although the precise magnitude of the gap varies across samples, tests, and ages, it is nearly always above 0.5 standard deviations and gaps approaching one full standard deviation are not uncommon.[1] Moreover, while the gap shrunk rapidly during the 1970s and 1980s, progress largely stopped among cohorts born between the early 1970s and the late 1980s (Neal 2006; Chay, Guryan, and Mazumder 2009). O'Neill (1990) and Neal and Johnson (1996) find that the test score gap has important implications for later economic outcomes, so slow progress in closing the test score gap suggests that economic disparities will persist for many decades to come.

Ethnographic evidence (Kozol 1992, Lareau 2003) suggests that families' material circumstances can account for much of the black-white gap. But this view has not been supported by statistical analyses of representative samples. Jencks and Phillips (1998) conclude that "[i]ncome inequality between blacks and whites appears to play some role in the test score gap, but it is quite small" (p. 9) and that "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). The gap has also proven to be surprisingly resilient to controls for other family characteristics. Hedges and Novell (1998), for example, find that differences in parental education and family income explain only about 30% of the black-white gap, though 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.

---

[1] See, e.g., the reviews by Neal (2006), Jencks and Phillips (1998), and Magnuson and Waldfogel (2008). Phillips et al. (1998) and Fryer and Levitt (2004, 2006) find smaller gaps for the very youngest children.

Fryer and Levitt (2004, 2006) are the most successful at explaining the gap via differences in observable characteristics. Controlling for a list of covariates ranging from the child's birth weight to the number of children's books in the home, they find no residual gap in kindergarten reading scores and only a small gap in kindergarten math scores. However, the residual gap grows with age, to nearly 0.4 standard deviations by the end of the 3rd grade.

In this paper we argue that important shortcomings in the way that income is measured have led the existing literature to dramatically understate the role of family income differences in accounts of the black-white test score gap and, therefore, to dramatically overstate the gap among children with the same family incomes. Studies of the conditional black-white gap typically control for the family's measured income in the year that the child was tested. As has long been recognized (Modigliani and Brumberg 1954; Friedman 1957), annual income is a poor proxy for a family's consumption and investment possibilities, and in any case it may be measured with substantial error in population surveys. To see the implications of this, suppose that test scores depend on permanent income but that a researcher controls only for annual income, a noisy proxy. Standard errors-in-variables results mean that the income coefficient will be attenuated relative to what would be obtained with the correct income measure.[2] As mean income is lower for blacks than for whites, this attenuation leads to overstatement of the black-white test score gap conditional on income.[3]

---

[2] We assume for the moment that annual income equals permanent income plus a i.i.d. error. Although this specific configuration is unlikely (Haider and Solon 2006), more general income processes produce similar results for the conditional black-white gap.
[3] Friedman (1957) makes exactly the same point about black-white differences in consumption.

In literatures where income is the explanatory variable of interest researchers often attempt to form better measures of permanent income.[4] But the insight that current income is an inadequate proxy for family resources 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 rely on other variables – like maternal education or socioeconomic status indices – to proxy for the family's long-run prospects (Fryer and Levitt 2004, 2006).[5]

We begin with an analysis of data from the Child Supplement to the National Longitudinal Study of Youth (CNLSY). We show that current income is at best a limited proxy for family resources, and that it explains only about half of the black-white gap in long-run incomes. Accordingly, in specifications for children's test scores at age 10 or 11, the conditional black-white gap falls by about 50 percent more (relative to the unconditional gap) when long-run average income is controlled than when an annual income measure is used. This result is robust to a variety of plausible deviations from the permanent income specification, including specifications that control in various ways for the time profile of the family's income as well as for its average level or that allow past or current income to matter more than does future income.

---

[4] Blau (1999), Kalil and Wightman (2011) and Mayer (1997) all explore achievement specifications that include medium- or long-run income averages, but none examine the impact of this on conditional black-white achievement gaps.

[5] In the sociological literature, researchers often control for wealth rather than income. We discuss these studies below (see note 9). Blau and Graham (1990) and Altonji and Doraszelski (2005) account for permanent income in their studies of black-white wealth gaps, as does Sanandaji (2009) in analyzing racial differences in mortgage approval.

One reason that researchers studying test score gaps do not control for long-run income measures is that the data requirements are onerous: Child scores are rarely available in the same data sets as the longitudinal family income records needed to construct income histories. We describe how instrumental variables (IV) techniques can be used to identify the black-white test score gap conditional on permanent income even when only annual income is observed in the test score sample, relying on an auxiliary data set with long income histories.

The IV strategy is also useful when test scores and income histories are available in the same sample. Under certain conditions – which appear to hold in the CNLSY – the IV estimate is less affected by measurement error in permanent income than is OLS. Accordingly, we find that the income coefficient is larger and the conditional black-white gap smaller in a specification where long-run income is instrumented with current income. In this specification, the conditional black-white gap is only 0.32 standard deviations, down from 0.56 without income controls, 0.43 when annual income is controlled, and 0.36 when long-run average income is controlled in an OLS specification. The IV estimate of the conditional black-white gap is little changed when we loosen the permanent income assumption by allowing current income to have an independent effect on current achievement.

We next turn to data on 5th graders from the Early Childhood Longitudinal Study (ECLS). We find that the black-white math score gap conditional on permanent income and a short list of family structure variables (e.g., mother's age) is only 0.18 standard deviations. By contrast, the gap without income controls is 0.62 standard deviations and the traditionally-estimated gap conditional on annual income is 0.38. We also reconsider Fryer and Levitt's (2006) analysis of 3rd grade math scores. When we add permanent income to the vector of

controls used by Fryer and Levitt we find that the remaining black-white gap in 3rd grade falls from 0.34 to 0.15 standard deviations.

Our analysis is purely descriptive: We do not attempt to distinguish the causal effect of family income from the effects of other characteristics (e.g., parental ability or "culture") that might be correlated with income. Our goal is simply to provide a more accurate description of the data than is possible with current income alone. Our results are thus only suggestive about the possible impact of income-focused interventions. Nevertheless, they call into doubt the frequent interpretation of the apparent robustness of the black-white gap as evidence that the gap is primarily due to black-white differences in characteristics such as genes (Herrnstein and Murray 1996), culture (Moynihan 1965), or parenting styles (Brooks-Gunn et al. 1996).

In Section II, we discuss the interpretation of regressions that control for family income in observational data. Section III discusses the effect of using a noisy income proxy and describes our approaches to identifying the black-white gap conditional on true family resources. In Section IV, we discuss the two data sets used in this paper, the CNLSY and the ECLS. In Section V, we present simple analyses of the dynamics of family income in the CNLSY data. Section VI presents results on black-white test score gaps. Section VII concludes.

II. The role of family resources in test score regressions

We model educational production as depending on exogenous family or student characteristics, which we label "ability," a, and on educational investments, e: $s = f(a, e)$, where s is the student's measured achievement. Investment is chosen by the family, which must allocate resources Y between investment and unitary consumption, c. For the moment, we ignore any dynamic aspects of the investment decision. The family's allocation must satisfy the

budget constraint $c + \pi e \leq Y$, where π represents the price of educational expenditures. Subject to this constraint, e is chosen to maximize $U(s, c; \gamma)$, where γ represents a preference parameter.[6]

Using the implicit function theorem we can write the chosen expenditure level as a function of ability, preferences, prices, and resources, $e=g(a, \gamma, \pi, Y)$. Substituting this into the educational production function $f(a, e)$ and linearizing, we obtain

$$s \approx \beta_0 + a\,\beta_a + \gamma\,\beta_\gamma + \pi\beta_\pi + Y\beta_Y. \tag{1}$$

Here, the effects of tastes, prices, and resources operate solely through investment choices: $\beta_\gamma = f_e{}^*g_\gamma$, $\beta_\pi = f_e{}^*g_\pi$, and $\beta_Y = f_e{}^*g_Y$. By contrast, ability has both direct and indirect effects: $\beta_a = f_a + f_e{}^*g_a$.

Equation (1) is not estimable, as a, π, and γ are not readily observed. However, it is useful in understanding the sources of black-white gaps in s. 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 (1), we can write the unconditional black-white test score gap as

$$\delta(s) \approx \delta(a)\,\beta_a + \delta(\gamma)\,\beta_\gamma + \delta(\pi)\beta_\pi + \delta(Y)\,\beta_Y. \tag{2}$$

There are thus four sources of gaps in mean test scores: Differences in ability distributions ($\delta(a)<0$), differences in preferences ($\delta(\gamma)<0$), differences in the price of educational investment ($\delta(\pi)<0$)[7], and differences in resources ($\delta(Y)<0$).

Next, consider the gap conditional on resources Y. For any variable X, let $\delta_Y(X) = E[X \mid b=1, Y] - E[X \mid b=0, Y]$ be the black-white gap in X conditional on Y. By (1), we have

$$\delta_Y(s) = \delta_Y(a)\beta_a + \delta_Y(\gamma)\beta_\gamma + \delta_Y(\pi)\beta_\pi \tag{3}$$

---

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

[7] This might arise, for example, if discrimination in the housing market makes access to high-quality schools more expensive for black than for white families.

Conditioning on Y thus eliminates one of the three terms from (2). Evidence that the black-white test score gap is largely robust to controls for Y – i.e., that $\delta_Y(s)$ is nearly as large as $\delta(s)$ – would therefore suggest that the gap is primarily attributable to black-white differences in ability, preferences, or prices rather than to the direct effects of family resources.

By contrast, a small $\delta_Y(s)$ could be consistent with a raw gap that derives primarily from the causal effect of resources on investments (i.e., from β_Y) or with a gap due primarily to ability, attitudes, or prices that are well proxied by income (i.e., $|\delta(a)\ \beta_a + \delta(\gamma)\ \beta_\gamma + \delta(\pi)\ \beta_\pi|$ is large but $|\delta_Y(a)\ \beta_a + \delta_Y(\gamma)\ \beta_\gamma + \delta_Y(\pi)\ \beta_\pi|$ is not). Absent a strategy for isolating variation in Y that is independent of a, γ, and π, these two explanations cannot be distinguished. We focus below on recovering the conditional gap in observational data, which has been the focus on many previous studies (including those cited above).[8]

III. Income proxies

A. What is the relevant measure of resources?

Assuming that family resources are linearly related to the omitted factors discussed in Section II, the conditional black-white gap $\delta_Y(s)$ can be estimated as the b coefficient in a regression of test scores on b and Y:

$$s_i = \theta_0 + b_i\ \theta_b + Y_i\ \theta_Y + \varepsilon_i. \tag{4}$$

We have not yet specified the resource measure, Y, however. Researchers estimating (4) typically use the family's income in the year of the test, sometimes supplementing this with family characteristics such as maternal education. This rules out any dynamic aspect of the educational production process, but dynamic considerations are clearly relevant. Educational

---

[8] Mayer (1997) and Dahl and Lochner (2010) attempt to isolate the causal effect of family resources, β_Y. Dahl and Lochner in particular find that the causal effect of income is quite large. We return to this study below.

investment decisions are made at different points in time, and the timing of investments may matter for test score impacts. If so, an appropriate specification should condition on the resource constraint that applies to the investment decision in each period prior to the date at which achievement is measured (Todd and Wolpin 2007).

The Permanent Income Hypothesis (PIH) potentially offers a way around the resulting complexity. If this hypothesis is correct – and if families know their future incomes and needs with certainty – then investment decisions at every date depend only on the family's total lifetime income, both past and future, and not on the time pattern of that income. This suggests that the family's permanent income is a sufficient statistic for the family resources that influence each investment decision.

But PIH may not hold exactly. In particular, if families are unable to borrow against their future incomes or are uncertain of their future incomes, investments may depend more strongly on past than on future income.[9] Alternatively, PIH may simply be an inaccurate model of the family decision process. For example, family decisions may be made on a cash-in-hand basis, without saving or borrowing across periods. In that case, investments in year t would depend only on income in year t, regardless of the past or future.

A cross-cutting consideration is measurement error in resources, however defined. As we show below, mismeasurement of resources biases the conditional gap $\delta_Y(X)$ toward the

---

[9] The sociological literature on test score gaps focuses on wealth, which better corresponds to resources available for investment than even average past income. See, e.g., Yeung and Conley (2008), Phillips et al. (1998), and Orr (2003). However, wealth is notoriously difficult to measure accurately. Moreover, families with the same wealth may have different future incomes. Indeed, if PIH holds, wealth (i.e., savings) will be negatively correlated with future income conditional on past income. We do not attempt to control for wealth here. Thus, the conditional-on-permanent-income test score gap that we estimate may partially reflect black-white differences in receipt of bequests (Oliver and Shapiro 1997; Blau and Graham 1990; Altonji and Doraszelski 2005).

9

unconditional gap $\delta(X)$, with greater bias the worse is the resource measure. Studies that compare self-reported family annual income to administrative data generally find reliabilities of the former in the range of 0.7 to 0.9 (Marquis, Marquis, and Polich, 1986), though some estimates are as low as 0.6 (Coder, 1992). Given this, longer-run averages may be more reliable proxies for true current income than is measured current income. But even over long time periods the measurement error in income is not fully averaged out: Mazumder (2001) finds that a 15-year income average has a reliability of only about 0.8 as a proxy for lifetime income.

Finally, recall that family resource controls serve both to absorb black-white differences in budget constraints and to proxy for differences in other family characteristics such as ability or attitudes. The appropriate resource measure for one purpose may not be the same as for the other. Permanent income is likely a better proxy for permanent family characteristics than is current-year income, even for cash-in-hand consumers.

In the analysis below, we focus on the family's permanent income as the appropriate resource measure. We also explore other options, however, with specifications that allow educational outcomes to depend instead on income to date, on the current year's income, or on the time pattern of the family's income.

B. The test score gap conditional on a noisy proxy

Let y_i represent a noisy proxy for the truly relevant resource measure Y_i. For example, Y might be permanent income and y current income. Importantly, if black families have lower resources on average than white families ($\delta(Y) < 0$), the black-white test score gap conditional on y will overstate the gap conditional on Y.

Let $\alpha = \text{cov}(y,Y)/\text{var}(Y)$ be the coefficient of a projection of y onto Y. If y is merely a noisy measure of Y, then $\alpha = 1$. Under more general income processes α may vary with the life

cycle (Haider and Solon, 2006).[10] We assume that $y - \alpha Y$ is uncorrelated with both b and ε, and that neither $V(Y \mid b)$ nor $V(y - \alpha Y \mid b)$ varies with b. With these assumptions, the conventional errors-in-variables formula can be used to relate the misspecified regression of s on b and y,

$$s_i = \theta_0' + b_i \theta_b' + y_{it} \theta_y' + \varepsilon_i', \tag{5}$$

to the correctly-specified regression (4).

Replacing the Y in regression (4) with y in (5) has two effects. First, it rescales resources by the multiplicative factor α. Thus, if $var(y-\alpha Y) = 0$, $\theta_y' = \theta_Y / \alpha$ and $\theta_b' = \theta_b$. Second, if $var(y-\alpha Y) > 0$, the resource effect is attenuated and – assuming that $\theta_Y > 0$ and $\delta(Y) < 0$ – the black coefficient is biased downward:

$$\theta_y' = (R_b / \alpha_t) * \theta_Y < \theta_Y / \alpha_t \text{ and} \tag{6}$$

$$\theta_b' = \theta_b + (1-R_b) \delta(Y)\theta_Y < \theta_b. \tag{7}$$

where

$$R_b \equiv \alpha^2 \frac{V[Y \mid b]}{V[y \mid b]} = 1 - \frac{V[y \mid Y,b]}{V[y \mid b]} \tag{8}$$

is the within-race reliability of y as a proxy for Y.[11]

Alternatively, θ_b' can be expressed as a weighted average of the black-white gap conditional on full resources, θ_b, and the unconditional gap, $\delta(s) = \theta_b + \delta(Y)\theta_Y$: $\theta_b' = R_b \theta_b + (1-R_b) \delta(s)$. Intuitively, black families on average have lower true resources than white families with the same measured resources, with the difference increasing in $(1-R_b)$. As a consequence, controlling only for the noisy resource measure will produce a conditional gap shaded toward the unconditional gap, with more shading the lower is R_b.

---

[10] See Appendix B for empirical evidence on this life cycle variation.
[11] Under our assumption that corr(y-αY, b)=0, R_b is weakly smaller than the traditional, unconditional reliability measure $R \equiv 1 - [var(y \mid Y) / var(y)]$, with equality only if $\delta(Y) = 0$.

C. *Avoiding bias*

The simplest way to avoid the bias due to a noisy resource proxy is to control for Y directly. But this is often infeasible. We thus explore methods for recovering the parameters of interest when Y is not observed in the test score sample. Our approach is based on equations (6) and (7), which can be rearranged to express the coefficients of interest in terms of the feasible coefficients, the reliability measure R_b, and the scaling factor α:

$$\theta_Y = \theta_Y' * (\alpha / R_b) = \theta_y'/\lambda_y \text{ and} \tag{9}$$

$$\theta_b = \theta_b' - (1-R_b)\, \delta(Y)\theta_Y = \theta_b' - \lambda_b\, \theta_Y. \tag{10}$$

where $\lambda_y = R_b / \alpha$ and $\lambda_b = (1-R_b)\delta(Y)$.

Notice that λ_y and λ_b are simply the coefficients from a regression of Y on y and b. Thus, θ_Y can be estimated as the ratio of the coefficient from a regression of s on y to the coefficient from a regression of Y on y, in each case controlling for b. This is instrumental variables (IV), or two-stage least squares (TSLS), using the noisy proxy y as an instrument for the true resource measure Y.[12] Usefully, the first stage coefficients λ_y and λ_b can be estimated from an auxiliary sample that lacks a test score measure. When this is done, the calculation described here is the two-sample two-stage least squares (TS2SLS) estimator examined by Inoue and Solon (2010).[13]

In some of the specifications below, we include covariates – e.g., the number of books in the home – that are not available in auxiliary sample. Appendix C describes a version of the

---

[12] An alternative computation uses R_b to create an alternative resource proxy $\tilde{y} \equiv \frac{1}{\alpha}(R_b y + (1 - R_b)E[y \mid b])$ that shrinks the observed proxy variable toward its race-specific mean. A regression of s on b and $\tilde{y}$ then estimates θ_b and θ_Y. If R_b is estimated as the y coefficient from a regression of Y on y and b, $\tilde{y}$ is the fitted value from this regression and the resulting two-stage least squares estimate is mechanically identical to the IV estimate described in the text.

[13] Inoue and Solon (2010) show that TS2SLS is not identical to the two-sample instrumental variables estimator considered by Angrist and Krueger (1992). Both our main approach and the equivalent shrinkage approach discussed in footnote 12 are implementations of TS2SLS.

TS2SLS estimator that we use in this case. The identifying assumption is that the covariates available only in the main sample are uncorrelated with y conditional on Y.

D. Measurement error in permanent income

In the NLSY data we are able to follow families only through the middle of the parents' careers. Thus, our broadest resource measure is $\bar{y}_{15}$, the log of the family's average income over a 15-year period. Insofar as investments depend on income outside this period, the errors-in-variables analysis above implies that both an OLS regression of s on $\bar{y}_{15}$ and b and an IV regression that instruments for $\bar{y}_{15}$ with y will overstate the magnitude of the black-white gap conditional on Y. If $y_t = Y + e_t$, the ratio of the bias in OLS to the (asymptotic) bias in IV equals var($\bar{y}_{15}$ | Y, b) / cov(y_t, $\bar{y}_{15}$ | Y, b). If the e_t are independent and identically distributed, the numerator and denominator of this expression both equal $\sigma_e^2 / 15$, so the two regressions should yield similar, and similarly biased, estimates. Under more general e_t processes, however, the asymptotic bias in the IV estimator may be larger or smaller than that in OLS.

In an empirically relevant case, the bias is smaller in IV than in OLS. Suppose that the e_t are independent but not identically distributed across t. Then the above ratio will be greater than one – the bias in the IV estimator will be smaller than that in OLS – if $V(e_t) < \frac{1}{15}\sum_\tau V(e_\tau)$.[14] Family income is less volatile in middle age than it is early in the life cycle, so an IV estimator using income from the former period as the instrument is less biased by measurement error in $\bar{y}_{15}$ than is OLS. As we discuss in Section VI, in our sample the "current" family income measure comes from a relatively low-noise age. Thus, we explore specifications that use y as an instrument for $\bar{y}_{15}$ even when $\bar{y}_{15}$ is observable in the test score sample.

---

[14] With non-independent e_t, the condition is $\text{cov}\left(e_{it}, \frac{1}{15}\sum_\tau e_{i\tau}\right) < \frac{1}{15}\sum_r \text{cov}\left(e_{ir}, \frac{1}{15}\sum_\tau e_{i\tau}\right)$.

IV. Data

Our analyses draw on two nationally representative samples. The first is the 1979 National Longitudinal Survey of Youth, a sample of over 12,000 teens and young adults in 1979 who 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. At each survey, respondents are asked detailed questions about their family incomes from various sources. 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 the scores available for (for example) 6-year-olds depend 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 we control for the age (in months) at which the exam was taken.[15] Scores on each test are normalized to mean zero and unit variance based on the CNLSY's 1968 norming sample.

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 1957 and 1964. 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.

---

[15] The testing protocol was not always followed perfectly. We use scores taken between ages 9.5 and 12.5, and exclude children with no scores in this three-year window.

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.

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. 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 income from an unmarried partner – available only in later waves – from the family income calculation.

We consider a variety of family resource measures. The first is the current income, the family's total income from all sources in the year in which a CNLSY child was tested. The second is a long-run income average, the simple mean of the real family income (in 2005 dollars) over the years in which the mother was aged 25 to 39. We refer to this as the family's permanent income. We also sometimes examine averages over shorter periods – e.g., four or six years prior to the CNLSY test, or over the entire period from the mother's age 25 until the test year.

In each survey year, roughly one-fifth of our sample has missing values in one or more of the income components. If every respondent missing information from any minor income component in any survey – e.g., someone missing food stamp benefit information in 2002 – were excluded from our permanent income calculation, we would have values for only 29% of CNLSY children with test score data. Moreover, even observations with complete data from each survey are missing income data for odd-numbered calendar years after 1994, when the NLSY switched to a biennial survey cycle.

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 (2010). This algorithm, described in Appendix A, allowed us to form a usable current income for 99% of CNLSY children (unweighted) and a permanent income measure for 94%, with missing values arising primarily when mothers permanently attrited from the NLSY sample before age 39.[16] Log current incomes average 10.72 (standard deviation 0.97), while the log permanent incomes average 10.77 (SD 0.70).

Our second sample is the Early Childhood Longitudinal Survey (ECLS) Kindergarten Cohort. This panel, the basis for several recent studies of the black-white test score gap (e.g., Fryer and Levitt 2004, 2006), follows a random sample of 21,000 students who were enrolled in kindergarten in the 1998-1999 school year. Our analysis focuses on students' math scores from the spring of 5th grade (and, in some analyses, from the spring of 3rd grade), as this permits a rough comparison to the similarly-aged CNLSY sample. We use scaled item response theory (IRT) scores, standardized to have mean zero and unit variance.[17]

The ECLS test scores are likely preferable to those available in the NLSY – the tests are more modern, and the testing regime more systematic. But the ECLS income measures are of much lower quality than those in the NLSY. Each wave of the ECLS contains a single income variable, the parent's report of the total income of all persons in the household, assigned to one of 13 bins. 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. We use the income reported in the spring of 5th grade as the current income for analyses of 5th grade test scores. We also construct a short-run

---

[16] Most respondents from the NLSY's military sample and economically disadvantaged white oversample were dropped from the panel relatively early. Thus, these subsamples represent only 0.1% of our main analysis sample. Our NLSY analyses all use custom longitudinal weights for children who are in the survey for any year 1979-2004.

[17] Strictly, the "5th grade" test is the one given in 2004 to students who were in kindergarten in 1999 or first grade in 2000; most but not all students were in 5th grade at the time. The IRT model is updated with each wave of the ECLS, producing changes in each prior score. Our analyses of both 3rd and 5th grade scores use scores from the 5th grade data release.

average income from the responses in the springs of Kindergarten, 1st, 3rd, and 5th grades. We set this to missing unless there are at least three non-missing values, two non-imputed.

We exclude from our analyses of both the CNLSY and ECLS any respondent who is not either black or non-Hispanic white. Tables 1A and 1B show summary statistics for the two data sets. The first column of each table conditions only on the availability of a math score. The second columns exclude families for which we are unable to construct the relevant income variables or our core demographic controls (age, gender, maternal age).[18] The third and fourth columns show statistics for the black and white subsamples.

V. Permanent and Current Income in the NLSY

Table 2 presents simple analyses of the relationships between race, permanent income, and annual income in the CNLSY. Our base sample is the same as that in Column 2 of Table 1A: Black and non-Hispanic white children with non-missing demographics and family income (current, lagged, and permanent). Columns 1 and 2 report simple bivariate regressions of log current and long-run average income, respectively, on an indicator for being black. Column 1 shows that black students' families have current log incomes 0.87 below those of white students' families, on average. This gap falls to 0.64 when we control for gender, child's age, mother's age, year, and birth order (Panel B). The raw gap in long-run average incomes (Column 2) is smaller, but most of the difference disappears when the maternal age control is added.

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

---

[18] In Table 1A, we exclude observations for which we were unable to measure or impute the family income two or four years before the test date.

indistinguishable from one and the R^2 just over 0.6.[19] The black coefficient is small and, when

controls are included, insignificantly different from zero, consistent with our assumption that the

transitory component of current income when children are age 10-11 is pure noise.

Column 4 reverses this regression, placing long-run average income on the left-hand side

and current income on the right. This is the first stage regression for our 2SLS analyses

presented below. Here, the current income coefficient (which corresponds to $\lambda_y = R_b / \alpha$) is just

above 0.5. The black coefficient λ_b is negative, -0.27, 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 long-run income gap without current income controls from Column 2.

Thus, test score regressions that control only for current income will dramatically understate – by

nearly half – the explanatory power of family income for the black-white test score gap.

Columns 5 and 6 explore alternative controls that shed light on the family income

dynamics that underlie our analysis. Column 5 adds to the specification in Column 4 a control

for the log of the average family income in the years prior to the child's test date (back to

mother's age 25).[20] Not surprisingly, this measure explains much of the variation in long-run

average income and reduces the remaining black-white gap by over 80%. However, the gap

remains statistically significant. Thus, insofar as educational investments depend on expectations

---

[19] The income coefficient in this regression estimates α. Note, however, that our sample is based
on the child's age rather than the mother's. Thus, if α is assumed to vary over the life cycle of
the mother (Haider and Solon, 2006), the regression estimates a weighted average with weights
corresponding roughly to the age distribution of mothers of 10-11 year olds in the CNLSY. See
the discussion in the appendix.

[20] We exclude a few observations where the child was tested before the mother was 25. In
roughly 1/3 of our sample, the child was tested after the mother turned 39, so the pre-test average
income is identical to our permanent income measure.

of future income, even controls for current and past incomes will not fully absorb racial resource differences relevant to current test scores.

Column 6 explores the degree to which the low current income coefficient and large conditional black-white gap in column 4 can be attributed to pure measurement error in current income. In this specification, we instrument for income in the year of the child's test with income two years later. As these measures are collected in different surveys, many sources of measurement error should be independent between them. Consistent with the measurement error hypothesis, the income coefficient is larger in this specification than in Column 4, and the remaining black-white gap shrinks by about two thirds. As in Column 5, however, it remains statistically significant, suggesting that pure measurement error corrections will reduce but not eliminate the overstatement of the conditional-on-resources black-white gap.

We have also explored life-cycle variation in the difference between current and permanent income. The right panel of Figure A1 in the appendix presents the standard deviation of y_{it} conditional on $\bar{y}_{15}$, as a function of the mother's age. This is high in the early 20s, as incomes are quite volatile at this point in the life cycle, then declines to a low point that is maintained from the late 20s through about 35 before rising again. The right panel of Figure A2, also in the appendix, shows what this implies for the transitory component of family incomes at different points in children's lives. The shocks are smallest at ages 7-12 and higher on either side of this point. Evidently, family income when a child is aged 10 or 11 has lower variance than does income at other ages. As discussed in Section III.D, this implies that an OLS regression of student achievement on $\bar{y}_{15}$ will be attenuated to a greater degree than is a 2SLS specification that instruments $\bar{y}_{15}$ with annual income in the year of the child's test.

VI. Results

A. Evidence from the CNLSY

Table 3 presents regressions for student scores on the PIAT math exam, given to members of the CNLSY sample at age 10 or 11. Column 1, which uses the maximal possible sample, shows that the raw black-white gap is 0.77 standard deviations. Column 2 (and the remainder of our analysis) restricts the sample to families for whom we observe enough information to compute a permanent income and annual incomes 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.[21] These controls bring the gap down to 0.56. Most of the change from Column 2 reflects between-race differences in the distribution of mother's age at the child's birth.

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 15-year 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.

---

[21] We have estimated all specifications without the controls, with similar results.

Column 8 presents the two-stage-least-squares specification discussed in Section III, using the log of current income as an instrument for the log of the long-run average income. As noted earlier, life-cycle variation in the transitory component of income means that the 2SLS estimates should be less attenuated by measurement error in long-run income than are those obtained by OLS. Indeed, we find that the income coefficient is larger and the black coefficient smaller (in magnitude) than in the corresponding OLS specification in Column 7.[22]

The exclusion restriction for the 2SLS estimator is that current income has no direct effect on student achievement, conditional on the family's permanent income. This restriction might fail for a variety of reasons, ranging from credit constraints to myopia to uncertainty about future income. To gauge the sensitivity of our results to violations of the assumption, Table 4 presents several specifications that allow in various ways for a direct effect of current income.[23]

We begin by returning to the OLS specifications controlling for current and long-run average income separately (Columns 4 and 7 of Table 3, respectively), repeated as Columns 1 and 2 of Table 4. Column 3 puts both income measures into the same specification. The long-run average income coefficient is over three times as large as that on current income, but the latter is nevertheless statistically significant. At first glance, this casts doubt on our 2SLS strategy. However, the exclusion restriction cannot be tested in this way – the explanation advanced earlier that both current and permanent income are measured with error would predict exactly the same result. Specifically, with $\text{cov}(y_{it} - Y, \bar{y}_{15} - Y) < \text{var}(\bar{y}_{15} - Y)$, as discussed

---

[22] One implication of our measurement error discussion is that a 2SLS specification that uses income from a younger or older age as the instrument would be attenuated to a greater degree. We find exactly this when we repeat the 2SLS specification using family income six years before the test date (when the child was aged 4 or 5) as an instrument: The income coefficient drops to 0.326, and the black coefficient increases in magnitude to -0.365.

[23] Caucutt and Lochner (2005) also examine the relationship between the timing of family income and student achievement in the CNLSY.

above, the probability limit of the current income coefficient in Column 3 would be positive even if current income had no effect conditional on permanent income.

Columns 4-6 present three 2SLS specifications that attempt to distinguish the two explanations. First, in Column 4, we include only current income, but instrument it with income four years prior. Insofar as measurement error is independent across observations four years apart, this should purge the effect of measurement error. The income coefficient is larger than in Column 1 but still notably smaller than in Column 2, suggesting that measurement error in current income alone cannot account for our results. Column 5 repeats the 2SLS specification from Table 3, this time using income two years *after* the test as the instrument. Obviously, this cannot have a direct effect on the student's test score; it nevertheless yields a very similar permanent income coefficient to our earlier estimate.

When we add current income as a control variable in this specification (Column 6), its coefficient is much reduced from Column 3 and is statistically insignificant. This is consistent with the measurement error explanation for the earlier result but is hard to reconcile with an explanation based on a causal effect of current income. More importantly, the black coefficient in this column is nearly identical to that from our earlier 2SLS specification. This is unsurprising, given the result from Table 2 that there is no black-white gap in current income once long-run income is controlled. Nevertheless, it suggests that any effects of current income that we miss in these specifications are unlikely to affect our main conclusions.

In Table 5, we explore the sensitivity of our results to another category of PIH violations. If young families are uncertain about future income or are credit constrained, early career income may be a more important determinant of educational investments than is realized future income. We thus explore specifications that allow income arriving before the test date to have a distinct

effect from income arriving later. Column 1 presents our baseline OLS specification (from Table 3, Column 7) for comparison. Column 2 replaces the long run average income control with the log average income in the years prior to the test (back to maternal age 25), as in Table 2. The income coefficient is somewhat reduced and the black coefficient somewhat larger, though the changes are very small. In Column 3 we control for both income measures simultaneously. The permanent income coefficient is much larger and the past income coefficient is statistically indistinguishable from zero, though positive. This provides at best weak evidence that past income matters more. Again, the more important result for our analysis is that the black coefficient is essentially unchanged across specifications. Evidently, an average of past and future income does a slightly better job of absorbing cross-racial differences than does past income alone, and the two averages entered separately do a better job still, but the differences are very small.

Finally, in Column 4 we discard both income averages in favor of separate controls for log family income in each year from (maternal) age 25-39.[24] We cannot reject the hypothesis that all of the income coefficients are the same, nor that the early age average equals the later age average. The black coefficient is slightly smaller in magnitude than in Column 1, but still larger than in our 2SLS estimates in Tables 3 and 4.

Taking the estimates in Table 3, 4, and 5 together, three things are clear. First, simply including annual income in a regression severely under-controls for differences in resources between black and white families. The black coefficient in our 2SLS specification (Column 8 of

---

[24] Lubotsky and Wittenberg (2006) recommend including all available proxies separately rather than choosing among them (but do not consider strategies like our IV estimate). We have also tried controlling for the family's long-run average income and for a family-specific age gradient in income. In this specification, the gradient coefficient is insignificant and the black coefficient is -0.357 (0.038).

Table 3) is only three-quarters as large as that in our OLS specification controlling for current income (Column 4 of Table 3). Stated somewhat differently, the inclusion of an annual income control explains just over half as much of the raw black-white gap (as in Table 3, Column 3) as is explained by permanent income. Even specifications that adjust for measurement error in current income indicate larger conditional gaps than do specifications that control for the long-run average. Second, we find only weak and statistically insignificant evidence that the time pattern of income predicts student achievement, conditional on the long-run average, and little sign that this time pattern provides much information about black-white test score gaps. Indeed, in both OLS and IV specifications in Tables 4 and 5, the black coefficient is remarkably stable to different ways of treating the time path of income. In contrast, income prior to the test does not fully absorb black-white differences in family circumstances that are relevant to children's test scores, though the explanatory value added by controlling for subsequent income is quite small. Thus, analyses that control for a sufficiently long income history can come quite close to the results obtained when both past and future incomes are controlled.

Table 6 presents estimates of our main specifications 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.[25] However, 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.

---

[25] Recall from Table 1A that the sample standard deviation of PPVT scores is much larger than those for PIAT scores, perhaps indicating a problem with the NLSY score norms. Nevertheless, the black-white gap on the PPVT is notably larger than on either PIAT component even when measured in within-sample z-score units.

B. Evidence from the ECLS

Table 7 presents estimates for students' 5th grade math scores in the ECLS. Column 1 shows that the raw black-white gap from the maximal possible sample is 0.85 standard deviations. This is shrinks to 0.78 when we restrict the sample to observations for which we have 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 quadratic controls for the mother's age at the child's birth. These are necessary for our two-sample analyses, as the CNLSY sample is only representative conditional on the mother's age. Maternal age 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 dramatically, 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 about one-third larger here, and the black-white gap shrinks to 0.34.

In Column 7, we present our TS2SLS specification that uses the CNLSY data to estimate the first-stage relationship between permanent income and the instrument, current income, and uses the ECLS data to identify the reduced-form relationship between current income and test scores (as in Column 5).[26] As before, the identifying assumption is that the transitory component of income is uncorrelated with achievement, conditional on race and our other controls. The income coefficient is over 50% larger than in Column 6 (and more than double that in Column

---

[26] Unfortunately, the age distributions of mothers in the CNLSY and the ECLS differ due to differences in the sampling schemes of the two surveys, and the first-stage coefficients may depend on the maternal age distribution (Haider and Solon, 2006). When we reweight the ECLS sample to match the maternal age distribution in the CNLSY sample, the raw black-white gap is somewhat smaller (0.62 vs. 0.78 standard deviations) but the pattern of coefficients across specifications is similar.

5). The conditional black-white gap is 0.18, less than a third of the raw gap and just over half of the gap controlling for the average income over the ECLS panel.

These estimates almost certainly undercorrect for the role of true permanent income. 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.[27] If so, the income coefficient in the TS2SLS specification remains attenuated, and the black coefficient somewhat negatively biased.

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 these controls may absorb a portion of the variation in permanent income conditional on current income, thus partly 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. We focus on two widely-available and commonly controlled variables that are plausibly good proxies for permanent family income, maternal education and the presence of a father.[28] Of course, these variables may have direct effects on student achievement.

The first panel of Table 8 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 current income, Column 3 uses an

---

[27] In the CNLSY, we construct a binned current income variable for use in the first stage, to correspond to the measure available in the ELCS. However, we do not adjust for the added precision presumably gained by the use of a full income module in the CNLSY survey rather than a single question in the ECLS.

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

average income over a longer period instead, and Column 4 presents estimates using our 2SLS (TS2SLS in Panel B) correction. Not surprisingly, the raw black-white gap is reduced by the inclusion of maternal education and father presence controls (compare Column 1 of Table 8 to Column 3 of Table 3 and Column 4 of Table 7). Less expected is that the specification that includes current income yields larger black-white gaps than in the analogous specifications without the new controls. Evidently, conditional on income black students have somewhat *better* family situations than whites. Or, put somewhat differently, mother's education and father's presence do not fully explain the black-white gap in family incomes. The pattern of results across Columns 2, 3, and 4 of Table 8 is similar to that seen earlier: Even when maternal education and family structure are controlled, a model with current family income overstates the conditional black-white gap by 23 (CNLSY) to 63 (ECLS) percent relative to what is obtained when long-run income is controlled via our 2SLS 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. FL (see also Fryer and Levitt 2004), showed that differences in covariates explained roughly the same absolute black-white gap across specifications for Kindergarten, 1st, and 3rd grade scores, but that the unexplained gap grew monotonically across these grades. Columns 1 and 2 of Table 9 report the FL estimates of the 3rd grade raw math score 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.[29] The raw gap is 0.88 standard deviations, and the inclusion of the FL controls reduces this to 0.38.

---

[29] FL include Hispanics and Asians in their sample, with dummy variables for each group. The coefficients on these dummies are not reported in Table 9.

Columns 3 and 4 reproduce Fryer and Levitt's analysis, restricting the sample to just blacks and non-Hispanic whites to correspond with the other estimates presented in this paper.[30] 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. The income coefficient is small but significant, while the black coefficient shrinks slightly but is generally similar to that seen in Column 6.

Because not all of the FL variables are available in the NLSY, we must use our hybrid TS2SLS estimator (see Appendix C) to adjust the FL results for permanent income. The key assumption of this estimator is that the transitory component of current income is uncorrelated with any of the ECLS-only control variables conditional on the covariates that are available in both samples. This is clearly false for the socioeconomic status index, as this is constructed from current family income. Columns 8 and 9 repeat the estimates from Columns 6 and 7 without this index. The specification without our family income control yields a slightly larger black-white gap, but that with a control for current income yields a notably smaller gap (and much larger income coefficient) than when the SES index is included.

Even with the mechanically related SES index excluded, the exclusion restriction may be incorrect. It would be violated, for example, if current income were correlated with the number

---

[30] Even when we include the other racial groups, we do not precisely reproduce Fryer and Levitt's sample or results. The most likely explanation 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. Other differences between our analysis and the FL specification are that we take control variables from the 3rd grade survey where possible, where FL appear to have used the kindergarten survey as the source of most covariates; we use 3rd grade cross-sectional weights in place of FL's longitudinal weights; and we present heteroskedasticity-robust standard errors where FL appear to report classical standard errors.

of children's books in the home 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 our estimator, in Column 10, we see that the long-run average income coefficient is more than double the current income coefficient in Column 9, while the black coefficient is only -0.15. This is just over half of the estimate from a specification with a current income control and much less than half of what is estimated without income controls at all (with or without the SES control). Evidently, even FL's rich specification is unable to effectively control for income differences between black and white families.

We have also reproduced Fryer and Levitt's (2004, 2006) analyses of test score gaps over time in the ECLS sample. Fryer and Levitt found that covariates explain much of the black-white gap in Kindergarten but that both raw and conditional gaps grow monotonically through 3rd grade. Our TS2SLS specification corroborates this result. Like Fryer and Levitt, we find that the gap explained by differences in observables is approximately stable across grades. The unexplained gap is smaller in the TS2SLS specification than in Fryer and Levitt's specification in each grade – indeed, in Kindergarten we find that black students earn _higher_ math scores than white students with similar observables, though the difference is insignificant – but as in their results it grows as students progress from Kindergarten through 5th grade.

VII. Discussion

Previous research has 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, these studies typically control only for family income in the year that the student is

tested, perhaps accompanied by additional covariates like maternal education. There is little theoretical justification for believing that current income is a sufficient control for the family resources that determine student achievement, and empirically both current income and human capital measures turn out to be very poor proxies for long-run measures of families' financial circumstances.

We describe a method for identifying the black-white test score gap conditional on the family's average lifetime income that can be used even when the data set containing student test scores does not itself permit accurate measurement of lifetime income. Our method 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 2009).

We find that the association between family permanent income and student achievement is roughly twice as strong as that between current income and achievement. In our preferred 2SLS and TS2SLS estimates, a 10% increase in family permanent income is associated with an increase in child math scores of 0.04 (CNLSY) to 0.07 (ECLS) standard deviations. These coefficients cannot be interpreted causally, as they reflect both the true causal effects of family resources and the confounding effects of other factors that are correlated with both income and economic outcomes. The most obvious omitted variables – e.g., parental ability – would tend to bias the income coefficient upward relative to the causal effect of family income. Our estimated income coefficients are much *smaller*, however, than the plausibly causal effects of family income estimated by Dahl and Lochner (2010).[31]

---

[31] Dahl and Lochner (2010) estimate that a permanent $1,000 increase in a family income causes test scores to rise by about 0.06 SDs. Using the median family income in their sample, around $18,000, this implies that a 10% increase in family income would lift scores by 0.11 SDs. Dahl

Understatement of the income coefficient produces overstatement of the black-white test score gap conditional on income. In both the CNLSY and ECLS samples, we find that conventional methods understate the share of the black-white test score gap that is attributable to family income differences by about half. Where the prior literature has indicated that relatively little of the gap can be attributed to family income, we find that family financial circumstances can explain 40 to 75% of the raw 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 already-small unexplained gap in their specification.

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. This is not the pattern that one would expect if income is merely proxying for noneconomic family factors. Thus, although our analysis is purely descriptive, it does offer some hope that improvements in black families' economic circumstances could, absent any other changes, lead to substantial closing of the black-white test score gap.

and Lochner speculate that the large effect that they estimate may reflect the low incomes of the disadvantaged families on which their estimates are based or a higher propensity to invest lump-sum EITC payments – which form the basis of their research design – than ordinary income.

References

Altonji, Joseph G. and Ulrich Doraszelski (2005). "The Role of Permanent Income and Demographics in Black/White Differences in Wealth," *Journal of Human Resources* 40(1), Winter, 1-30.

Angrist, Joshua D. and Alan B. Krueger (1992). "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association* 87 (418), June, 328-336.

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 John W. Graham (1990). "Black-White Differences in Wealth and Asset Composition," *Quarterly Journal of Economics* 105(2), May, 321-339.

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.

Caucutt, Elizabeth M. and Lance Lochner (2005). "Borrowing Constraints on Families with Young Children," *Proceedings*, Federal Reserve Bank of Cleveland, 49-48.

Charles, Kerwin, Eric Hurst, and Nick Roussanov (2009). "Conspicuous Consumption and Race," *Quarterly Journal of Economics* 124(2), May, 425-467.

Chay, Kenneth Y., Jonathan Guryan, and Bhashkar Mazumder (2009). "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth." National Bureau of Economic Research Working Paper No. 15078.

Coder, J.F. (1992). "Using Administrative Record Information to Evaluate the Quality of the Income Data Collected in the Survey of Income and Program Participation," *Proceedings of Statistics Canada Symposium 92: Design and Analysis of Longitudinal Surveys.* Ottowa Ontario: Statistics Canada.

Dahl, Gordon and Lance Lochner (2010). "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." Unpublished manuscript, Nov. 19.

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.

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.

Inoue, Atsushi and Gary Solon (2010). "Two-Sample Instrumental Variables Estimators," *Review of Economics and Statistics* 92, August, 557-561.

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.

Kalil, Ariel and Patrick Wightman (2011). "Parental Job Loss and Children's Educational Attainment in Black and White Middle-Class Families," *Social Science Quarterly* 92(1), March, 57-78.

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

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

Lubotsky, Darren and Martin Wittenberg (2006). "Interpretation of Regressions with Multiple Proxies," *Review of Economics and Statistics* 88(3), August, 549-562.

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.

Marquis, Kent H., M. Susan Marquis, and J. Michael Polich (1986). "Response Bias in Sensitive Topic Surveys," *Journal of the American Statistical Association* 81(394), June, 381-389.

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 Working Paper 2001-24.

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.

Neal, Derek and William Johnson (1996). "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy* 104 (5), October, 869-895.

O'Neill, J. (1990). "The Role of Human Capital in Earnings Differences between Black and White Men," *Journal Of Economic Perspectives*, 4 (4), Fall, 25-45.

Orr, Amy J. (2003). "Black-White Differences in Achievement: The Importance of Wealth," *Sociology of Education* 76(4), October, 281-304.

Oliver, Melvin L. and Shapiro, Thomas M. (1997). Black Wealth/White Wealth: A New Perspective on Racial Inequality. Routledge: New York.

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.

Sanandaji, Tino (2009). "Reversion to the Racial Mean and Mortgage Discrimination." Institutet För Näringslivforskning working paper 811.

Todd, Petra E. and Kenneth I. Wolpin (2007). "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps," *Journal of Human Capital* 1 (1), Winter, 91-136.

Yeung, W. Jean and Dalton Conley (2008). "Black-White Achievement Gap and Family Wealth," *Child Development* 79(2), March/April, 303-324.

Appendix A: Data

In this appendix, we describe the imputation procedure that we use to fill in missing values in the NLSY income variables. 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 is too much information missing to permit this, we use coarser imputation procedures, though we use these only to construct permanent income; we exclude observations for which current income needed to be imputed this way. Our procedure is based loosely on that used by Dahl and Lochner (2010), 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 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 convert family incomes to 2005 dollars and censor the annual values at $3,373 (the 5th percentile in our sample). We form our permanent income by averaging these censored real incomes over the years when the mother is aged 25-39.

We exclude from our samples observations for which our current or permanent income measures require excessive imputation. First, we drop individuals who attrit from the survey before age 39, for whom we would have to extrapolate family income to years outside of the range for which we have actual values. Second, we exclude individuals for whom we have to interpolate the family income aggregate for any survey year or for more than two of the non-survey years used in the permanent income calculation. Finally, we drop individuals for whom employment status in the year that the child took the test must be imputed.

Our analysis of the NLSY uses custom sampling weights generated for the universe of CNLSY respondents who appear in any survey between 1986 and 2006. In the ECLS, we use weights appropriate for the grade-5 cross-section of children (C6CW0).

Appendix B: Income process

Haider and Solon (2006) assume that $y_{it} = \alpha_t Y_i + e_{it}$. Letting t index maternal ages, we estimate α_t by regressing current income y_{it} on long-run average income $\bar{y}_i$ for different values of t. The regressions are estimated on our main NLSY sample, estimating $\bar{y}_i$ as the average of family income from age 25 to 39 and allowing t to vary over the same range. The α_t coefficients, along with 95% confidence intervals, are shown in the left panel of Figure A1, while the right panel shows the root mean squared errors from these regressions. α_t begins low but rises to about 1.1 by the early 30s and stays relatively constant through the end of the 30s. The transitory component, $y_{it} - \alpha_t \bar{y}_i$, is most variable among the oldest and youngest women, with relatively little variation for women in their early 30s.

Our main analysis focuses on 10-11 year old children, whose mothers vary in age. As Table 1A indicates, the average mother in our NLSY sample gives birth in her mid 20s but there is substantial variation around this average. Figure A2 shows the average of α_t and of the standard deviation of the transitory component as functions of the child's age. α rises monotonically, while the transitory component is less variable for 9-year-old children than for older or younger children. Vertical lines in the figure show the average age of CNLSY children at the date of testing. This is near the minimum of the transitory variation curve.

Appendix C: The hybrid TS2SLS estimator

In some of our specifications using the ECLS data, we include covariates that are available in the ECLS but not in the NLSY. This requires adapting the instrumental variables estimator, as the first-stage regression of permanent income (Y) on current income (y) and the controls cannot be estimated in the NLSY data. Our proposed estimator blends elements of two-sample two stage least squares (TS2SLS) and two-sample IV (TSIV).

To describe the proposed estimator, it is useful to convert to matrix representations. Let W represent non-income covariates, including b, a constant, and any other controls that are to be added to (5). Let $Z = [y \ W]$ and $X = [Y \ W]$. Suppose that Z and s are observed in sample 1 (the "test score sample") and that Z and X are observed in sample 2 (the "auxiliary sample," which may or may not be the same as sample 1), and let subscripts denote the sample in which a variable is measured. The estimand is the coefficient vector from equation (4), θ.

The two-sample two-stage least squares (TS2SLS) estimator,

$$\hat{\theta}_{TS2SLS} = \left(\left(Z_2'Z_2 \right)^{-1} Z_2'X_2 \right)^{-1} \left(Z_1'Z_1 \right)^{-1} Z_1's_1.$$ (C1)

exploits the fact that Z is a valid instrument for X. If the four second moment matrices in (C1), when scaled by the appropriate sample sizes N_1 and N_2, consistently estimate the corresponding population moments, the TS2SLS estimator identifies θ. (In a contrast from traditional IV applications, the OLS estimator $\hat{\theta}_{OLS} = (X'X)^{-1} X's$ would be consistent for θ. However, it is infeasible if X and s are not observed in the same sample.)

To introduce covariates V that are available in sample 1 but not in sample 2, redefine Z = [y W V] and X = [Y W V]. We require consistent estimates of E[Z'X] and E[Z's]. The latter can be obtained solely from sample 1. But the former,

$$E[Z'X] = \begin{pmatrix} E[y'Y] & E[y'W] & E[y'V] \\ E[W'Y] & E[W'W] & E[W'V] \\ E[V'Y] & E[V'W] & E[V'V] \end{pmatrix},$$ (C2)

cannot. With one additional assumption samples 1 and 2 can be combined to estimate each element of (C2). Specifically, let $\lambda = [\lambda_Y \ \lambda_W']'$ be the coefficients of a linear projection of y on Y and W. λ can be estimated from sample 2. We assume that $E[V'(y - Y\lambda_Y - W\lambda_W)]=0$; that is, that V is uncorrelated with the noise component of the resource proxy conditional on the control variables W. With this assumption, E[V'y] = E[V'Y]λ_Y + E[V'W] λ_W, so E[V'Y] = (E[V'y] − E[V'W]λ_W)λ_Y^{-1}. This permits a hybrid estimator for E[Z'X]:

$$n_2^{-1} \tilde{Z}'\tilde{X} = n_2^{-1} \begin{pmatrix} y_2'Y_2 & y_2'W_2 & \frac{n_2}{n_1} y_1'V_1 \\ W_2'Y_2 & W_2'W_2 & \frac{n_2}{n_1} W_1'V_1 \\ \frac{n_2}{n_1}\left(V_1'y_1 - V_1'W_1 \hat{\lambda}_{W,2} \right) \hat{\lambda}_{p,2}^{-1} & \frac{n_2}{n_1} V_1'W_1 & \frac{n_2}{n_1} V_1'V_1 \end{pmatrix} \xrightarrow{p} E[Z'X].$$ (C3)

We form a corresponding Z'Z matrix:

$$n_2^{-1} \tilde{Z}'\tilde{Z} = n_2^{-1} \begin{pmatrix} y_2'y_2 & y_2'W_2 & \frac{n_2}{n_1} y_1'V_1 \\ W_2'y_2 & W_2'W_2 & \frac{n_2}{n_1} W_1'V_1 \\ \frac{n_2}{n_1} V_1'y_1 & \frac{n_2}{n_1} V_1'W_1 & \frac{n_2}{n_1} V_1'V_1 \end{pmatrix} \xrightarrow{p} E[Z'Z].$$ (C4)

We use these to form a hybrid two-sample 2SLS estimator:

$$\hat{\theta}_{HTS2LS} = \left(\left(\tilde{Z}'\tilde{Z} \right)^{-1} \tilde{Z}'\tilde{X} \right)^{-1} \left(Z_1'Z_1 \right)^{-1} Z_1's_1.$$ (C5)

This is consistent for θ.[32]

---

[32] We could replace $n_2^{-1}\tilde{Z}'\tilde{Z}$ with any consistent estimator for E[Z'Z], including $n_1^{-1}Z_1'Z_1$. The choice of $n_2^{-1}\tilde{Z}'\tilde{Z}$ follows from Inoue and Solon's (2010) intuition for the superiority of TS2SLS to the two-sample IV estimator $\hat{\theta}_{TSIV} = \dfrac{n_2}{n_1}\left(Z_2'X_2\right)^{-1}Z_1's_1$, as it adjusts for sampling differences between samples 1 and 2 that appear in $n_2^{-1}\tilde{Z}'\tilde{X}$. Consistent with this, Monte Carlo simulations suggest that the estimator based on $n_2^{-1}\tilde{Z}'\tilde{Z}$ performs better than one based on $n_1^{-1}Z_1'Z_1$.

Figure A1. Income process parameters, by mother's age
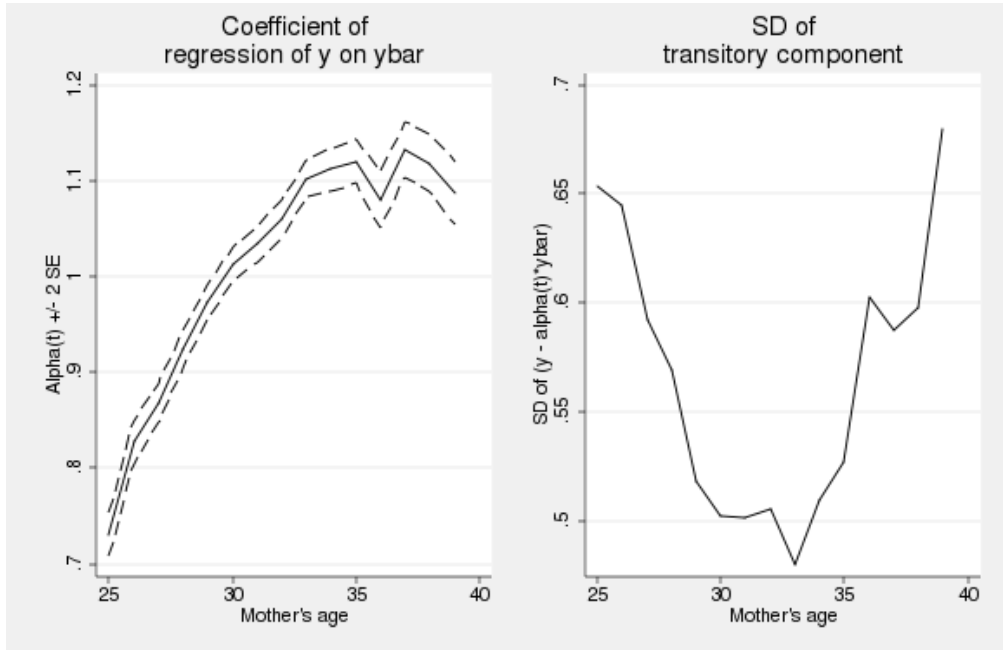
Figure A2. Average income process parameters, by child's age
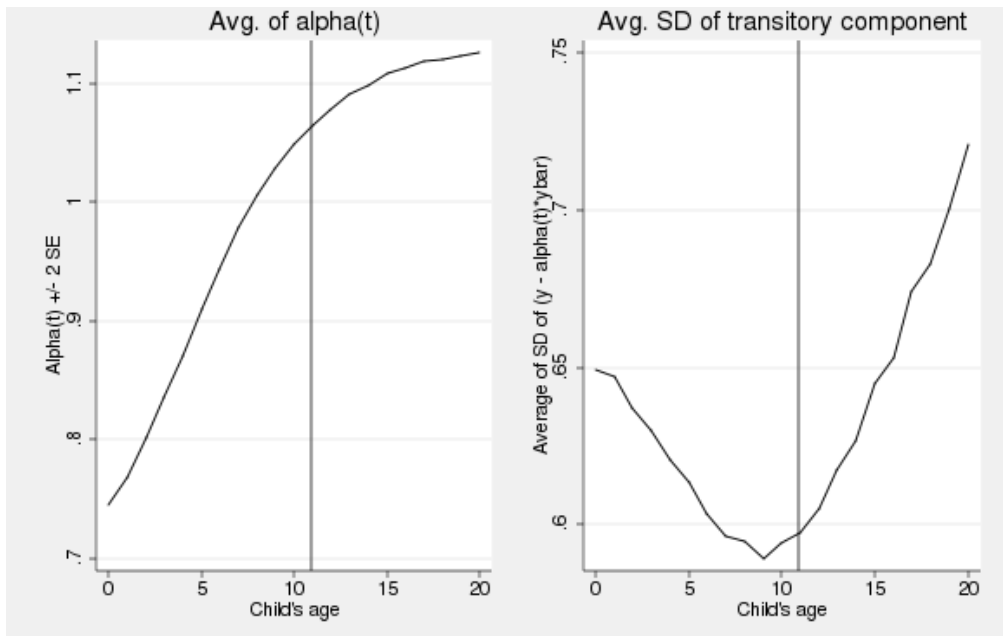
Table 1A. NLSY summary statistics

	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]
Missing both PIAT scores	0.02	0.02	0.04	0.02
Missing PPVT score	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]	[0.70]	[0.70]	[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]	[5.20]	[4.99]
Observations	5440	4966	2061	2905

Notes: Table displays means; standard deviations are in brackets. Calculations 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, lagged, and permanent income measures. Mean test scores are nonzero because the NLSY uses a 1968 norming sample.

Table 1B. ECLS summary statistics

	Population (1)	Sample (2)	Black (3)	White (4)
Math score	0.06	0.18	-0.46	0.31
	[0.99]	[0.96]	[0.92]	[0.91]
Reading score	0.09	0.19	-0.39	0.31
	[0.98]	[0.96]	[0.95]	[0.92]
Female	0.48	0.48	0.48	0.48
Age at test (months)	133.5	133.4	132.8	133.5
	[4.4]	[4.2]	[4.1]	[4.3]
ln(income)	10.80	10.89	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
	[2.29]	[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
	[0.45]	[0.43]	[0.49]	[0.38]
Observations	7742	6143	777	5366

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 whom 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)	Ln(long run average income)	Ln(current income)	Ln(long run average income)	Ln(long run average income)	Ln(long run average income)
	OLS	OLS	OLS	OLS	OLS	2SLS w y(t+2)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. No controls						
Black	-0.867	-0.760	-0.055	-0.301	-0.035	-0.116
	(0.039)	(0.032)	(0.028)	(0.024)	(0.013)	(0.026)
Ln(long run average income)			1.068			
			(0.024)			
Ln(current income)				0.529	0.132	0.751
				(0.014)	(0.010)	(0.020)
Ln(past income)					0.763	
					(0.017)	
R2	0.132	0.191	0.623	0.648	0.900	
Panel B. With controls						
Black	-0.642	-0.599	-0.022	-0.271	-0.046	-0.119
	(0.040)	(0.033)	(0.026)	(0.022)	(0.013)	(0.024)
Ln(long run average income)			1.035			
			(0.026)			
Ln(current income)				0.511	0.133	0.755
				(0.015)	(0.010)	(0.023)
Ln(past income)					0.786	
					(0.016)	
R2	0.216	0.271	0.631	0.657	0.906	

Notes: Sample is children in the CNLSY sample for whom tests were administered at age 10 or 11 and for whom contemporaneous, lagged, and permanent family income could be constructed. Column 5 controls for the mean income from the mother's age 25 to one year before the test date. N=4,966 for columns 1-4, N=4,910 for column 5, and N=4,547 for column 6. 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, are in parentheses.

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

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	OLS (7)	2SLS (8)
Black	-0.767	-0.758	-0.557	-0.426	-0.410	-0.404	-0.362	-0.318
	(0.033)	(0.034)	(0.035)	(0.037)	(0.037)	(0.037)	(0.038)	(0.041)
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	0.400
							(0.027)	(0.037)
Controls?	n	n	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 2-8. Sample in columns 2-8 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. Specifications allowing for direct effects of current income

Method	OLS	OLS	OLS	2SLS	2SLS	2SLS
Endogenous				y	ybar	ybar
Instrument				y(t-4)	y(t+2)	y(t+2)
	(1)	(2)	(3)	(4)	(5)	(6)
					2SLS	2SLS
Black	-0.426	-0.362	-0.360	-0.376	-0.319	-0.323
	(0.037)	(0.038)	(0.038)	(0.039)	(0.042)	(0.043)
Ln(current income)	0.205		0.079	0.283		0.026
	(0.019)		(0.026)	(0.031)		(0.047)
Ln(long run avg. income)		0.327	0.246		0.380	0.347
		(0.027)	(0.037)		(0.037)	(0.081)
N	4,942	4,942	4,942	4,942	4,524	4,524
R2	0.21	0.22	0.22			

Notes: Sample for all columns excludes children with missing family income variables. Columns 5-6 also exclude children for whom income two years after the test date is unavailable. SEs, clustered on the mother, in parentheses. See notes to Table 2 for control variables and weights.

Table 5. Past versus future income income

	(1)	(2)	(3)	(4)
Black	-0.362	-0.369	-0.359	-0.344
	(0.038)	(0.038)	(0.038)	(0.038)
Ln(long run avg. income)	0.327		0.255	
	(0.027)		(0.066)	
Ln(avg. income age 25 - test date)		0.316	0.077	
		(0.027)	(0.064)	
Ln(income) at each age 25-39, entered separately	n	n	n	y
N	4,942	4,930	4,930	4,942
R2	0.22	0.21	0.22	0.22

Notes: All specifications estimated by OLS. Sample for all columns excludes children with missing family income variables. Columns 2 and 3 also exclude those whose mothers were under age 25 at the test date. All specifications include baseline controls from Table 2, Panel B; column 4 also includes fifteen separate controls for income in each year between maternal age 25 and 39. SEs, clustered on the mother, in parentheses. See notes to Table 2 for control variables and weights.

Table 6. Sensitivity of black-white gaps on three NLSY test scores to alternative income controls

Method	OLS	OLS	OLS	2SLS
Income measure	None	None	Current	Long-run avg.
	(1)	(2)	(3)	(4)
Panel A: PIAT Math				
Black	-0.758	-0.557	-0.426	-0.318
	(0.034)	(0.035)	(0.037)	(0.041)
Ln(income) coefficient			0.205	0.400
			(0.019)	(0.037)
Controls	n	y	y	y
R2	0.10	0.18	0.21	
Panel B: PIAT Reading Composite (0.5*reading comprehension + 0.5*reading recognition)				
Black	-0.625	-0.464	-0.338	-0.232
	(0.033)	(0.035)	(0.037)	(0.041)
Ln(income) coefficient			0.197	0.386
			(0.019)	(0.037)
Controls	n	y	y	y
R2	0.08	0.15	0.19	
Panel C: PPVT				
Black	-1.278	-1.059	-0.909	-0.787
	(0.048)	(0.047)	(0.049)	(0.054)
Ln(income) coefficient			0.235	0.453
			(0.027)	(0.051)
Controls	n	y	y	y
R2	0.17	0.22	0.24	

Notes: See notes to Tables 2 and 3. Samples exclude families with missing current or long-run average income. N = 4,942, 4,864, and 4,623 in Panels A, B, and C, respectively. SEs, clustered on the mother, in parentheses.

Table 7. Sensitivity of black-white gaps on ECLS 5th grade math scores to alternative income controls

Sample	Full	Analysis					
Method	OLS	OLS	OLS	OLS	OLS	OLS	TS2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Black	-0.846	-0.775	-0.779	-0.623	-0.383	-0.343	-0.175
	(0.052)	(0.066)	(0.065)	(0.071)	(0.067)	(0.070)	(0.079)
Ln(current income)					0.334		
					(0.030)		
Ln(short-run avg. income)						0.434	
						(0.035)	
Ln(long-run avg. income)							0.698
							(0.066)
Gender, age	n	n	y	y	y	y	y
Mother's age at birth	n	n	n	y	y	y	y
R2	0.13	0.09	0.10	0.14	0.23	0.24	

Notes: Full sample is children in the ECLS sample for whom tests were administered in the spring 5th grade survey. N = 7,742. Analysis sample in Columns 2-7 excludes children for whom family income, average income, or any controls are unavailable: N = 6,143. Specifications in Columns 3-7 include controls for gender and a quadratic in the child's age at test date. Specifications in Columns 4-7 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 8. With controls for maternal education and the presence of a father figure

Method	OLS	OLS	OLS	2SLS / TS2SLS
Income measure	None	Current	Sample avg.	Long-run avg.
	(1)	(2)	(3)	(4)
Panel A: NLSY (PIAT Math)				
Black	-0.508	-0.457	-0.407	-0.371
	(0.035)	(0.037)	(0.038)	(0.043)
Ln(income)		0.144	0.238	0.322
		(0.023)	(0.032)	(0.051)
Mother's education	0.095	0.075	0.067	0.058
	(0.008)	(0.008)	(0.008)	(0.010)
Father present	0.123	0.001	0.005	-0.038
	(0.036)	(0.040)	(0.039)	(0.043)
N	4,926	4,926	4,926	4,926
Panel B: ECLS (Math)				
Black	-0.469	-0.396	-0.362	-0.244
	(0.070)	(0.068)	(0.071)	(0.078)
Ln(income)		0.225	0.295	0.550
		(0.032)	(0.036)	(0.082)
Mother's education	0.134	0.101	0.096	0.068
	(0.009)	(0.009)	(0.009)	(0.011)
Father present	0.231	0.042	0.062	-0.035
	(0.061)	(0.064)	(0.061)	(0.072)
N	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 7, respectively. The additional controls are mother's education (in 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 the NLSY) are in parentheses. Regressions in each panel use sampling weights.

45

Table 9. Analysis of ECLS 3rd grade math scores

	Fryer and Levitt (2006)		Full sample		Income subsample					
			OLS	OLS	OLS	OLS	OLS	OLS	OLS	TS2SLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black	-0.882	-0.382	-0.824	-0.396	-0.776	-0.344	-0.324	-0.371	-0.286	-0.147
	(0.031)	(0.033)	(0.031)	(0.033)	(0.038)	(0.039)	(0.039)	(0.041)	(0.040)	(0.051)
Ln(current income)							0.076		0.265	
							(0.022)		(0.017)	
Ln(permanent income)										0.556
										(0.046)
Socioeconomic status index		0.288		0.371		0.382	0.326			
		(0.015)		(0.018)		(0.020)	(0.026)			
Age (months)		0.019		0.020		0.022	0.022	0.020	0.022	0.010
		(0.002)		(0.003)		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Birth weight (oz)		0.003		0.004		0.004	0.004	0.005	0.004	0.005
		(0.000)		(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female		-0.175		-0.139		-0.164	-0.161	-0.163	-0.154	-0.135
		(0.018)		(0.022)		(0.024)	(0.024)	(0.025)	(0.025)	(0.029)
# of children's books		0.006		0.002		0.003	0.003	0.005	0.003	0.003
		(0.001)		(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
# of books squared (*1000)		-0.020		-0.006		-0.008	-0.007	-0.014	-0.008	-0.007
		(0.003)		(0.004)		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Mother over 30 at first birth		0.083		0.097		0.088	0.091	0.210	0.157	0.111
		(0.024)		(0.031)		(0.033)	(0.033)	(0.034)	(0.034)	(0.036)
Mother a teen at first birth		-0.132		-0.163		-0.155	-0.153	-0.246	-0.190	-0.128
		(0.025)		(0.033)		(0.037)	(0.037)	(0.039)	(0.038)	(0.039)
Mother receives WIC benefits		-0.208		-0.180		-0.162	-0.141	-0.379	-0.195	-0.025
		(0.024)		(0.030)		(0.033)	(0.034)	(0.032)	(0.034)	(0.045)
Constant	0.275	-1.576		-2.417	0.300	-2.654	-3.515	-2.576	-5.633	-7.751
	(0.012)	(0.168)		(0.307)	(0.015)	(0.339)	(0.428)	(0.349)	(0.398)	(0.569)
N	11,201	11,201	9,934	9,934	8,076	8,076	8,076	8,076	8,076	8,076
N (black or white)	7,908	7,908	9,934	9,934	8,076	8,076	8,076	8,076	8,076	8,076
R2	0.12	0.26	0.12	0.28	0.09	0.27	0.27	0.21	0.25	

Notes: Columns 3 and 4 use our full sample, which excludes children who attrited before the 5th grade interview. Columns 5-11 further exclude children with missing family income in the 3rd grade wave. Specifications in columns 2, 4, and 6-10 include as controls indicators for missing birthweight, children's books, mother's age at first birth, socioeconomic status (except in columns 8-10), and WIC receipt. Columns 3-10 use the ECLS 3rd grade cross-sectional weights. Robust SEs in parentheses. Column 10 uses the proposed adaptation of the TS2SLS estimator; see text for details.