Title
The NELS Curve: Replicating The Bell Curve

Authors
Levine, David I.
Painter, Gary

Publication Date
1998-02-04

The NELS Curve:

Replicating *The Bell Curve* Analyses with the

National Education Longitudinal Survey

by

David I. Levine
Haas School of Business
University of California
Berkeley CA 94720-1900
(510)642-1697
levine@haas.berkeley.edu

and

Gary Painter
School of Public Administration
University of Southern California
Los Angeles, CA 90089-0041
(213) 740-8754
gpainter@usc.edu

Abstract: This study uses the National Educational Longitudinal Survey of 1988 (NELS) to replicate both the analysis in *The Bell Curve* and that of several of its previous replications. We examine the relative importance of test scores and family background in predicting dropping out of high school, starting college, arrests, and out-of-wedlock fertility. Our results relax several arbitrary assumptions made in *The Bell Curve*. We strongly reject *The Bell Curve*'s conclusion that family background is almost always less important than test scores in predicting outcomes. In addition, our analysis casts doubt on some of *The Bell Curve*'s claims concerning reverse discrimination in education

Acknowledgment: We appreciate useful comments from Ken Chay, Claude Fischer, Paul Gertler, Sam Lucas, Matt Spiegel, and Kim Voss.

All programs used in this project are available from the second author. (Data are available from the National Center for Education Statistics.)

Institute of Industrial Relations
Working Paper Series
Working Paper No. 66

The Bell Curve is perhaps the most reviewed book in the social sciences. Even more of an honor (among scientists), its analyses are also among the most replicated (Currie and Thomas, 1997; Cawley, Heckman, et al., 1997; Dickens, Kane and Schultz, 1996; Hout, et al., 1996; Korenman and Winship, 1996). Until now, these replications have relied on the dataset analyzed by the authors of *The Bell Curve*, the National Longitudinal Survey of Youth (NLSY).

This study replicates both the original analysis in *The Bell Curve* and that of several of the previous replications using a new dataset, the National Educational Longitudinal Survey of 1988 (NELS) and its three follow-ups through 1994. This survey covers youth from approximately ages 14 to 20 (eighth grade to sophomore year of college). It permits testing many of the predictions of *The Bell Curve* concerning how test scores and family background affect educational outcomes such as dropping out of high school and attending college, as well as other outcomes such as teen pregnancy and arrests.

The analysis here proceeds in two phases. First, we follow Herrnstein and Murray's (H&M) methods in *The Bell Curve* rather closely. We replicate H&M's finding that test scores often being more important than their narrow index of socioeconomic status in predicting youths' outcomes, although the socioeconomic status index is relatively more important in our dataset. We then relax many of the arbitrary assumptions H&M make, as have other researchers who have replicated their results analyzing the NLSY. As these other replications have found, test scores remain important predictors of outcomes even with rich controls for family background. At the same time, and also as other replications have found, H&M's assignment of primary role to test scores is not robust to including richer measures of family background. In contrast, we find that the skills measured by standard cognitive tests are often only a fourth to a tenth as important as our measures of family background in predicting outcomes.

We also replicate H&M's analysis of black-white gaps in education, fertility out of wedlock, and arrests. H&M find that gaps in test scores explain (in a statistical sense) more than 100% of the gap in education. They interpret the resulting higher education of blacks than whites with similar test scores as evidence of reverse discrimination. We find that differences in family background are roughly as important as differences in test scores, and that the evidence is much less clear concerning the importance of reverse discrimination, differences in "culture," differences in skills and attitudes unmeasured by the tests, or test bias.

Method

Following H&M, this paper first analyzes a sample of white respondents to compare the relative importance of test scores and family background for predicting youths' outcomes. It then analyzes the relative importance of these two sets of factors in "explaining" (in a statistical sense) the racial gap in many of the outcomes.

Test scores "versus" family background

Following Herrnstein and Murray, we use an integrated framework for analyzing the determinants of a wide array of outcomes: high school dropout, entering college, unmarried teen pregnancy, and arrests. (The youth of the NELS respondents precludes an analysis of labor market outcomes such as wages.) The basic method *The Bell Curve*'s empirical section was a "horse race" between an index of test scores known as the AFQT and a measure of family background in predicting important life events. H&M assert that test scores explain most of the explainable variance of outcomes.

In their discussion of public policy, they provide evidence that large and long-term improvement in test scores are difficult to achieve. (Fischer, et al., 1996, summarize the evidence somewhat differently.) Thus, Herrnstein and Murray conclude that public policy interventions are unlikely to be successful at increasing the living standards of the disadvantaged.

As noted in the introduction to this Symposium, these horse races have two important problems when used to evaluate policy. First, test scores are not solely a measure of almost-immutable general cognitive ability, and the correlation between test scores and outcomes such as college attendance may not be solely due to how well scores measure cognitive ability.

For example, test scores are partly a function of family background. Endogenous test scores are particularly likely to be a problem in the NLSY, where many of the highest-scoring respondents did not complete the cognitive test until they were in college, while many of the lowest-scoring respondents had already dropped out of high school (Fischer, et al., 1996). Hence, education partially caused test scores in the NLSY, as well as being caused by them. The result is that correlations of test scores with outcomes such as completing high school can be misleading. Respondents to the NELS all took the first cognitive achievement test in eighth grade, diminishing this problem.

Although the NELS solves the problem of the test being given after differing amounts of schooling, the tests given in the NELS are designed to be school achievement tests in reading and mathematics. Cognitive tests such as those in the NLSY and NELS measure a combination of at least five factors: the trait of general cognitive ability; the quality of schooling and other influences on educational achievement; skill at taking tests; test bias; and other dimensions of the person's personality or upbringing such as comfort with taking tests and willingness to work hard at the meaningless task of taking a multiple choice test for researchers. A debate continues on the relative mix of these three factors for each test. (Fischer, et al. [1996] and H&M review this debate, coming to rather different conclusions.) The tests in the NELS, although not too different from the reading and mathematics sections of the AFQT in the NLSY, were specifically designed to measure what students have learned in school. As such, these cognitive tests should be more influenced by school quality than a test specifically designed to measure general cognitive ability.

When test scores are used to predict college attendance, they are particularly likely to over-state the importance of cognitive ability. Many colleges use tests very similar to those in the NELS to evaluate applicants (for example, the SAT). If people of similar cognitive ability vary in their comfort taking tests, then those who are most comfortable will typically have higher scores on both the NELS tests and college admissions tests. Thus, the correlation between the NELS test scores and college attendance may partially capture comfort with or motivation on tests, thereby over-stating the importance of cognitive ability.

The second problem with H&M's analysis occurs when they imply that if test scores are the most important predictors of outcomes, then policy interventions are not a good idea. In fact, for policy purposes what matters is the cost-effectiveness of an intervention, not the relative contribution a variable makes in predicting outcomes. Thus, if test scores predict ten percent of the variance of an outcome and school quality predicts only 1 percent, that result does not imply that we should abandon attempts to improve schools. Instead, potential improvements in school should be evaluated by comparing their costs and benefits.

With these cautions in mind, it remains interesting to identify the relative importance of test scores and family background in determining youths' outcomes. In posing a horse race between test scores and family background, H&M asked the wrong question. We are somewhat uncomfortable continuing to analyze this wrong question. At the same time, H&M published a widely-read set of answers that others have found to be misleading. Moreover, given that public policies may be made based on their analysis, it is important to examine whether H&M or their critics are correct in analyzing the relative importance of test scores and family background. As an additional justification for this study, H&M's results are easy to remember, even if not robust. It is important that if people are only going to remember a little bit about what determines success (cognitive ability vs. family background) that they remember the correct summary of the evidence (Dickens, et al., 1997). Our final justification is that by examining a different dataset that was collected a decade after the NLSY, we provide a check on the robustness of both sets of analyses.

The method of *The Bell Curve* is straightforward. Outcome variables were regressed (often using logistic regression) against a measure of test scores and an index of family background. (Sometimes other control variables such as age were included.) The index of family background was constructed as a weighted average of family income, the larger of the parents' education, and the larger of the parents' occupational status. The relative importance of test scores vs. the index of family background was measured by simulating how a one standard deviation increase in each affected the predicted probability of each outcome, holding the other predictor constant at its mean.

As a number of authors have pointed out, this "horse-race" procedure has two important limitations (Hout, et al., 1996; Dickens et al, 1996; Heckman, 1995; Korenman and Sanders, 1996). First, it restricts the coefficients of each component of the index of family background to have an identical effect. Thus, a one standard deviation increase in parental education is constrained to have the same effect on the outcome as a one standard deviation increase in family income. The assumption of equal effect has no theoretical basis. Furthermore, when the assumption is tested empirically, it is almost always rejected.

The second problem with this index is that it omits a number of family characteristics that generations of social science research have identified as important predictors of lifetime outcomes, including whether both parents are present, birth order, and region. Casting the net a bit wider, the technique omits measures of neighborhood and school characteristics that may also be important predictors of youth's success.

Thus, this study begins by replicating the analysis of *The Bell Curve*, utilizing H&M's restricted measure of family background. The study then replicates other research that has extended *The Bell Curve* analysis. That is, the regressions below "unbundle" the elements of the family background index and include additional measures of family and neighborhood background. Finally, the study adds a novel estimator to the methods used in the past: a within-school estimator that controls for all elements of the school's characteristics. (This method also captures important aspects of family and neighborhood characteristics.)

Because test scores and family background are correlated, no single metric of each set of variables' "importance" exists. Comparing the standardized coefficients of test scores and of family background, as done by H&M, is appropriate when there are only two variables, and neither variable is causally prior to the other. Unfortunately, neither of these assumptions is met in the analysis of family background and test scores.

Fortunately, we can bound the relative importance of each set of factors by examining their relative contribution to R^2. (Because we primarily use logistic regressions, we analyze a pseudo-R^2. To simplify the text we often refer to R^2, but all logistic analyses actually use pseudo-R^2.) An upper bound on a group of variable's importance is the R^2 of the regression when the group is entered alone. A lower bound on a group of variable's importance is the increment to R^2 when it is added after the other variables are entered first.[1]

The pseudo-R^2 for test scores entered alone may overstate the importance of test scores because family background is causally prior to test scores. Consider the case where all of the portion of test scores that is correlated with family background is caused by family background. In this case, the portion of the explained variance that is shared by both test scores and family background is actually due to family background. (In our data set, family background explains 27 percent of the variance of test scores among the white students.) In this setting, the appropriate measure of the role of the abilities measured by test scores is the incremental R^2 of test scores, after family background has already been entered into the regression.

[1] The pseudo-R^2 is defined as $(1 - L_1/L_0)$, where L_1 is the likelihood function with all the included variables, and L_0 is the likelihood function with only the constant. For the linear model, this formula yields the same results as the traditional R^2. In the logistic model, it rescales the likelihood function from zero (fits as well as the constant-only model) to 1 (perfect prediction; that is, log-likelihood=0) We follow H&M in relying on the pseudo-R^2 but we acknowledge that it is not precisely correct to use increments to pseudo-R^2 to measure "importance" of a set of

Entering family background first and examining the incremental R^2 of test scores yields an upper bound on the importance of family background. H&M argue that the correlation between child and parental cognitive ability is partially genetically based. In this case, parental genes for high cognitive ability may cause both good family background and, by inheritance, a youth's high test scores. The technique of entering family background first attributes these inherited abilities to family background. In fact, under this hypothesis, good parental genes directly increase children's inherited ability, and only incidentally lead to good family background. Thus, the upper and lower bounds are valid regardless of one's prior on the causal links between genes, family background, and test scores. Finer statements require more information than social scientists currently have.

Because social scientists traditionally compare individual coefficients, Fischer, et al. (1996) and Dickens, et al. (1997) create indices that sum up the effects of a bundle of family characteristics. Given the complicated causal paths between parental genetic endowments, family background, and test scores, it is safest to pay attention to results that are robust to both the upper and lower bounds provided by the changes in (psuedo-)R2. Nevertheless, social scientists are used to head-to-head comparisons in a regression framework. Thus, we computed how a one standard deviation "change" in family background affected each outcome and compared it to the effect of a one standard deviation change in test scores.

We used the method of Dickens, et al. (1997). We first estimated the logistic equation predicting outcomes as a function of family background and test scores. For each person, we multiplied the logit coefficient vector by that individual's family background to create a family background index. (That is, the family background index is a weighted average of people's family background, with the logit coefficients providing the weights.) The standard deviation of these scores was then used to represent a "one standard deviation increase in family background." The standard logit formula for the effect of a one unit change in an explanatory variable then gave us the change in probability of an outcome for someone one standard deviation above average in family background compared to someone with mean family background. (Specifically, dP/dX = m*(1-m)*s.d.(family background index), where m = the mean of the outcome variable. The dP/dX calculated with this formula is not be directly comparable to the odds ratio presented by

---

variables. As discussed below, our results are consistent enough that alternative measures of "importance" would not change them substantively.

the logit regressions because it evaluates a discrete change, while the logit examines continuous changes.)

Interactions: It is traditional in the literature on cognitive ability and family background to treat these two complicated phenomena as distinct. In fact, many outcomes may depend on an interaction of the two. For example, graduating high school may be almost a sure thing for a youth having high test scores or from prosperous families -- only the combination of lacking both puts the youth at risk. Conversely, entering graduate school may be especially likely when students have both high cognitive ability and prosperous families, while either one alone has little effect. Thus, we will test for interactions between family background and performance on cognitive tests.

The within-school estimator: A contribution of this paper is the ability to control for many features of the respondent's junior high school, neighborhood, and high school. This feature is possible because the dataset has multiple respondents (typically 25) for each junior high school. The method is to construct a within-school estimator that evaluates how youth with test scores higher or lower than their school average (or family backgrounds more or less privileged than their school average) achieve compared to the school average. Because we are estimating logistic equations, we use the conditional logit procedure to derive consistent within-school estimators (McFadden, 1974). The estimate is similar to a fixed effect estimator using panel data. (Because our outcomes are discrete, the ordinary least squares fixed effect estimator is not consistent [Hsiao, 1986].)

Controlling for average school quality with the within-school estimator eliminates the bias due to omitted school characteristics that are common for all students within the junior high school. This estimator also controls for characteristics of the family, high school, and neighborhood that respondents within a junior high school share.

Our within-school estimator controls for junior high school, although during most of our sample period the respondents were in high school Controlling for junior high schools is useful because junior high schools are usually smaller than high schools, implying that this measure captures more aspects of the respondent's neighborhood. In addition, the choice of junior high school may be less likely to be influenced by career plans or other outcomes we study. For example, over two percent of our sample changed schools to attend a magnet high school; many magnet schools have a career orientation.

Working in the other direction, somewhat more junior high school students attend private schools than do high school students (16.7% vs. 11.6%); the gap was smaller for non-religious private schools most likely to be endogenous (4.1% vs. 3.7%).[2] Controls for junior high school also omit important differences between the high schools that students from the same junior high may attend. Moreover, the within-school estimator does not control for all aspects of a students experience. For example, many schools put students in different tracks that lead to very different experiences. It is possible, although not clear, that family background and race influence the track students are placed on, even controlling for test scores (Gamoran, 1989; Lucas and Gamoran, 1991).

A further consequence of the NELS data having roughly 25 students per school is that the errors in logistic regressions are correlated among students. Thus, computed standard errors are biased downward. As discussed below, the conditional logit and logit results are typically quite similar. Moreover, our primary results have coefficient estimates that are many times their standard errors. Together, these results suggest this bias does not affect the interpretation of our results.

What Explains Racial Differences in Outcomes?

For most measures of social outcomes, blacks fare worse than whites. For example, blacks are more likely to drop out of high school, less likely to go to college, and more likely to bear a child out of wedlock. A crucial question is understanding the source of these differences. This question can be phrased as a statistical decomposition: how much of the difference in outcomes can be predicted by different characteristics of blacks and whites, how much of the difference can be predicted by different returns to the characteristics of blacks and whites, and how much remains unexplained even if blacks and whites had identical characteristics and returns (that is, how much is due to differences in the intercept [Oaxaca and Ransom, 1994]. As discussed below, this decomposition is a statistical exercise, not necessarily causal.

---

[2] The figures compare eighth and twelfth grades. Private non-religious schools includes all those not classified as religious.

In principle, the estimated black-white gap in an outcome is the difference in mean residuals for respondents of each race. As is usual with residuals, they represent our ignorance. In spite of this feature, social scientists and social commentators frequently interpret residuals. For example, many scholars have attributed the black-white wage gap to discrimination. Going to the other extreme, Juhn, Murphy, and Pierce (1991) attribute the wage gap to lower unmeasured human capital for blacks.

In some settings, racial gaps in outcomes favor the minority. The problem of interpreting the residual remains. For example, Mason (1996) attributes high black education levels (given their family background) as evidence of *high* levels of unmeasured human capital.[3] Mason's interpretation is consistent with generations of commentators on the high education of (for example) Jewish youth: some aspect of culture, energy, or skills not captured by standard measures lead to high education (conditional on observable measures).[4]

Conversely, Herrnstein and Murray find that blacks receive more education than whites with similar test scores. They interpret the high black education as evidence of reverse discrimination (pp. 319-322). Because including test scores in the equation explains (more than) 100% of the racial gaps in education, H&M conclude that gaps in family background have no room to explain Black's lower educational attainment.

Unfortunately, both H&M and Mason are interpreting residuals; neither provide direct evidence for their hypothesized interpretation. In fact, both sets of results are consistent with reverse discrimination, with test bias, or with unmeasured differences in effort, culture, or skills.

In our analysis we are unable to measure test bias or reverse discrimination directly, but we do have unusually good measures of test scores and of family background. Thus, we estimate equations for each outcome using a full set of interactions of each independent variable with a dummy variable for race:

---

[3] Haveman and Wolfe (1994: 170-173) and Maxwell (1994) also find that differences in family characteristics completely explain (in a statistical sense) differences between black and white rates of dropping out of high school.

[4] High unobserved human capital conditional on test scores is conceptually similar to and empirically identical to having a biased test. H&M cite several reviews of the evidence and conclude that no such bias exists (p. 281). When they later find evidence consistent with test bias, they interpret it as reverse discrimination. In contrast, Fischer, et al. (1996), and Dickens, et al., (1997) read the previous evidence on test bias as less conclusive.

$$\text{outcome} = b_0 + \mathbf{b_1 * family\ background} + b_2 * \text{test scores}$$
$$+ c_0 * \text{black} + \mathbf{c_1 * family\ background} * \text{black} + c_2 * \text{test scores} * \text{black}$$

Assuming identical returns to characteristics ($c_1 = c_2 = 0$), the coefficient on the Black indicator (c_0) would measure racial differences in outcomes given family background and test scores. This is the method followed in most studies of racial differences. Our more flexible decomposition permits us to test for differences in returns (c_1 and c_2 not equal to zero).

This method uses a first-order Taylor series to decompose the racial differences in outcomes into the shares due to differences in family backgrounds (b_1 * [mean family background of whites - mean family backgrounds of blacks]), differences in test scores (b_2 * [mean test scores of whites - mean test scores of blacks]), different returns to family backgrounds (c_1 * mean family background of whites), different returns to test scores (c_2 * mean test scores of whites), and different intercepts (c_0).

This formulation uses white coefficients and mean characteristics as the baseline. As a robustness check, we also perform the complementary decomposition using black coefficients and black mean characteristics as the baseline (Neumark, 1988).

For the purposes of these decompositions we estimate the outcome equations using ordinary least squares (OLS). The estimated coefficients from the OLS were very similar to the changes in probabilities we estimated with an identical probit equation. We relied on OLS estimation instead of probit or logit because the OLS decomposition is easier to understand.

These decompositions ignore the causal links between test scores, race, and family background. The raw gap in test scores between blacks and whites is .75 standard deviations. Controlling for the indicators of family background in this data set lowers this gap to .41 standard deviations. More complete measures of family background might further narrow the unexplained gap. The regression analysis we perform does not acknowledge the causal priority of family background, but puts test scores and family background on a level playing field.

Similarly, there is substantial evidence that blacks live in worse neighborhoods in large part due to discrimination in housing markets (e.g., Kingsley and Turner, 1993; Farley and Frey, 1992). Our method attributes some of the effects of racially-motivated housing discrimination as the effect of disadvantaged neighborhood -- again, a more complete causal model would increase the role of race.

The within-school estimator: It is possible that the decompositions of educational outcomes are largely due to differences in the quality of students' schools or neighborhoods.

Blacks are disproportionately concentrated in disadvantaged neighborhoods and attend less desirable schools. Many analysts predict that disadvantaged neighborhoods will disadvantage students. Thus, controlling for school characteristics, especially with use of a within-junior-high-school estimator, should reduce the black disadvantage in education and other outcomes.

At the same time, it may be that on average blacks attend less demanding high schools. In the NELS data, for example, if we examine blacks and whites with identical grade point averages, the blacks have on average test scores 0.68 standard deviation lower. It is possible that a student who would fail at a rigorous high school would successfully complete a high school degree in a less rigorous school. If lower rigor increases completion at (on average) less rigorous mostly minority schools, then blacks will have lower dropout rates than similar whites. In addition, to the extent that selective colleges use class rank, such colleges will give an advantage to students from less demanding schools -- that is, on average, to blacks. Thus, the combination of high school segregation by race and use of class rank in admissions can increase Black college attendance at selective colleges.

Data

The National Education Longitudinal Study of 1988 (NELS) is sponsored by the National Center for Education Statistics and carried out by the Bureau of the Census.[5] NELS is designed to provide trend data about critical transitions experienced by young people as they develop, attend school, and embark on their careers. The base-year (1988) survey was a multifaceted study questionnaire with four cognitive tests, and questionnaires for students, teachers, parents, and the school.

Sampling was first conducted at the school level and then at the student level within schools. The data were drawn from a nationally representative sample of 1,000 schools (800 public schools and 200 private schools, including parochial institutions). Within this school

[5] This information is drawn from the National Center for Education Statistics, 1994 and 1995.

sample, 25,000 eighth grade students were selected at random. The three follow-ups revisited the same sample of students in 1990, 1992, and 1994; that is, when most of the respondents were in the tenth grade, in the twelfth grade, and roughly two years after high school graduation.

Following H&M, we first study a sample of non-Hispanic whites. The second section compares non-Hispanic whites with Blacks.

The cognitive / achievement tests:[6] In addition to the student questionnaire, students completed a series of cognitive tests covering reading comprehension, mathematics, science, and history/citizenship/geography. "There is a reasonable consensus that [general cognitive ability] can be measured by summing across tests of several specific aptitudes, usually verbal and quantitative" (Chatman and O'Reilly, 1994). Thus, we make use of combined scores on the reading and mathematics tests.

The reading comprehension test included 21 questions. It contained five short reading passages or pairs of passages, with three to five questions about the content of each. Questions encompassed understanding the meaning of words in context, identifying figures of speech, interpreting the author's perspective, and evaluating the passage as a whole. The questions were drawn from various testing services including the National Assessment for Educational Progress (NAEP), the second international mathematics study, and the Educational Testing Service (Rock and Pollach, 1995).

The mathematics test had 40 questions. Test items included word problems, graphs, equations, quantitative comparisons, and geometric figures. Some questions could be answered by simple application of skills or knowledge, others required the student to demonstrate a more advanced level of comprehension and/or problem solving.

We analyze the equally-weighted average of the standardized reading and mathematics scores. For the small number of test takers (fewer than 1 percent) who had only a reading or a mathematics score but not both, the composite is based on the single available score.

The NELS design team created the composite score to be a "simple, overall... measure of cognitive ability to use as a control variable for cross-sectional analysis of data" (NELS, 1994). Unfortunately, the composite score is an imperfect measure of "cognitive ability." The test items

[6] These tests are described in more detail in Rock and Pollach, 1995.

are similar to those on a typical school achievement test; that is, items are designed to measure what most 14 year olds have been exposed to in classes. Although classroom learning is related to cognitive ability, it is affected by many other attributes of the family, teachers, school and neighborhood. Some of those additional factors are controlled for in the regressions below, but others remain omitted. Thus, the regressions below attribute to test scores the advantages both of high cognitive ability and of the other advantages of the youth's background that promote high levels of learning school material by age 14. (In spite of these cautions, H&M argue that related tests such as the SAT are good proxies for general cognitive ability. In addition, other researchers have used related achievement tests such as the GMAT as proxies for general cognitive ability [e.g., Chatman and O'Reilly, 1994].)

The tests have high reliabilities. The reliability of each subscore (measured as the 1 minus the ratio of the average measurement error variance to the total variance) was greater than .80, and often near .90 (Rock and Pollach, 1995: 67). We use the sum of the reading and math subscores, further increasing reliability. Unlike most tests, the NELS battery was also constructed to have similar reliabilities for different sex, racial and ethnic groups (Rock and Pollach, 1995: 4).

Socioeconomic Status Index and Rich Family Background: We use two measures of family background, a narrow one replicating of Herrnstein and Murray's socioeconomic status (SES) index, and a broader measure.

First we replicated H&M's SES index that combined parental occupational status, parental education, and family income into a single index. To do this, they converted each measure into z-scores with mean zero and standard deviation equal to one. They then took the maximum of the father's and the mother's centile of occupational status, using Duncan's measure of status. They coded "not in labor force" as the lowest status. For education, they took the father's and the mother's z-score on education. For the third component, they took the z-score of the log of family income, coding a zero income as -4. In each case, they replaced missing values with the mean value of zero. Finally, the H&M SES Index is the standardized sum of the four z-scores for maternal and paternal education, parental occupational status, and family income.

To construct a richer measure of family background, we relaxed a number of the arbitrary assumptions that underlie the H&M index. First, we did not constrain the coefficients on the several components to be equal; instead, we entered parental education, parental occupational status, and family income separately. We also had no prior that the effects would be linear, and that only the higher of the two parents' occupational status and education mattered. Thus, we entered the squared z-score of family income, and both parents' occupational status and education. For predicting dropouts and starting college, we included a measure of whether either parent dropped out or started college, respectively. Because students with missing values are likely to be missing parents, we included dummies for important missing values.

Following other replications of *The Bell Curve*, we included a number of variables that are potential indicators of advantaged family background: region, rural vs. urban vs. suburban, foreign language spoken in home, whether the mother or father foreign born, number of siblings, whether two parents are present in the home, and whether the home has a library card, magazines, and many books.

To supplement this fairly standard list, we drew from a wide range of measures that prior research suggested indicated advantages or disadvantages for youth. Thus, we included measures from all four of the NELS questionnaires for characteristics of the parents, teachers, students, and school.

From the parental questionnaire, we took indicators for whether the family was one of five religions, and any of five levels of religious observance. We controlled for whether the youth was in a single-parent home in eighth grade, his or her parents were divorced or separated while the youth was in high school, and whether a single parent became married. We constructed a variable indicating whether at least one parent was working full time (or, in two-parent households, if both were working part-time). We created this measure of "full-time equivalent worker" twice, once when the youth was in eighth grade and then again four years later. We controlled for whether at least one parent had been a teen when the youth was born. (We were unable to control for whether they were married at the time of the youth's birth.) As a catch-all to capture unmeasured aspects of the family and its support for education, we controlled for whether an older sibling had dropped out of high school by the time the respondent was in eighth grade. (This measure is a simple approximation to the within-family estimator. A within-family estimator controls implicitly for many characteristics of the family [Korenman and Winship, 1996].)

We were interested in measuring parent's involvement in the youth's life and education. Numerous commentators have suggested that discipline in the household, particularly concerning limits on TV, will increase student's academic achievement. Thus, we constructed simple indices of whether the family had rules concerning TV time, whether the family had other rules, and whether parents helped with homework more than once a week. As an indirect measure, we included a dummy variable for whether the child had participated in clubs such as Boy or Girl Scouts during elementary school. We also created a dummy equal to one if the parent belonged to a parent-teacher association or related organization, or volunteered at school.

To measure the characteristics of the students' classmates, we included dummy variables for whether the high school had less than 25 percent, or between 25 and 49 percent single parent households. We also included a dummy variable for whether the proportion minority was less than 40 percent.

From the teacher questionnaire we created an index of how dangerous the teacher reports the high school as being, and an index of whether the teacher reports drugs are a problem in schools. Finally, we included the student's report of whether teaching was good at the school.

Because this list is rather long, we reran the analyses with a narrower list of family background variables that are closer to standard descriptions of "advantage" or "disadvantage." These results are discussed below.

Interactions: We test for the presence of interactions by interacting our summary test score measure with the summary socioeconomic status index developed by H&M. This simple specification is parsimonious--future research can examine additional interactions.

Outcomes: We study a number of outcomes for each student. As in *The Bell Curve*, these include permanently dropping out of high school (that is, dropouts who do not receive a GED and do not start college), having a child out of wedlock, and self-reported arrests. H&M analyze whether respondents ever received a college degree. Because the NELS sample is only about age 20 in 1994, we study those who have started college.

A number of educational outcomes occur sequentially. Thus, only youth who complete one hurdle are eligible to try for the next more difficult one. We run a set of logistic regression on successively smaller samples: all sample members can complete high school, while only high school graduates can enter college. Finally, among those attending college, we examine the choice of two or four-year institutions. The advantage of running these as logits (as opposed to calculating expected years of education) is that the estimated effects of each independent variable can differ for different educational transitions. For example, a characteristic such as family income might matter more for the decision to go to college than for predicting who drops out of high school.

We analyze whether the respondent claims to ever have been arrested, as well as reports of ever having gone to juvenile hall. In addition to self-reports, H&M also analyze whether the respondent was ever interviewed in prison. Both miss substantial amounts of crime, whether due to under-reporting for the NELS, or lack of conviction or not being in jail during the interview week for the NLSY. There is no obvious reason to believe these differences will bias the results.

Summary statistics: Means and standard deviations are presented in Table 1. Six percent of the sample permanently drop out of high school, 78 percent have begun college, and 12 percent of the young women report having become pregnant out of wedlock.

Results of the Horse Race

Table 2 summarizes a number of regressions that compare the relative importance of family background and test scores in predicting youths' outcomes. The first rows use specifications similar to those of H&M, while the bottom rows analyze a richer measure of family background. Due to the large number of outcomes and specifications, the results are presented here in summary form. Complete regression results for predicting dropouts are in the Appendix, and results for other variables are available from the second author. Because our sample is of thousands of students and because the stratified sampling (multiple students per school) can lead to underestimated standard errors, we highlight as less convincing those results that are not statistically significant at the 1% level.

Dropouts: Test scores also have a large statistically significant effect on predicting which respondents drop out of high school. The pseudo-R^2 is 16%, and a youth with 1 standard deviation above-average test scores is one fifth as likely as the mean scoring respondent to drop out. The effects of a one standard deviation increase in test scores is similar to that of a one standard deviation increase in Herrnstein and Murray's socioeconomic status index when the two are entered simultaneously. This result is different from H&M, who find that the standardized coefficient on test scores is much larger. The difference may be due to differences in measures between the NLSY and the NELS, or it may be due to the fact that the NELS follows a decade later and SES has risen (relatively) in importance. If H&M's SES index causally precedes test scores, then SES is a somewhat more important contributor to pseudo-R^2: its R^2 alone is 14%, and the incremental R^2 of test scores is 8%.

A richer measure of family background greatly enhances its role. Rich family background and school characteristics have a pseudo-R^2 of 30% The incremental R^2 of test scores over rich family background is a modest 4%. At the same time, the point estimate of test scores is still large (though much smaller than with test scores entered alone): a high-scoring student is two-thirds as likely to drop out as a mid-scoring student. Conditional logit (within-school) estimates are similar to those with rich family background.

When we use the method that permits a head-to-head competition, family background remains more important than test scores. Specifically, a one standard deviation decrease in family background raises the odds of dropping out by 6.8 percentage points, while a one standard deviation decrease in test scores raises the odds of dropping out by 4.4 percentage points.

Starting college: The results for starting college closely parallel those for dropping out of high school. Test scores have a large statistically significant effect on predicting which respondents attend college by age 20. The pseudo-R^2 of test scores alone is 13%, and a youth with 1 standard deviation above-average test scores is three times as likely as the mean scoring respondent to begin college. The effects of a one standard deviation change in test scores is somewhat larger than the effect of Herrnstein and Murray's socioeconomic status index when the two measures are entered together. If the SES index causally precedes test scores, then SES is a more important contributors to pseudo-R^2: it's R^2 alone is 13%, and the incremental R^2 of test scores is 7%.

A richer measure of family background greatly enhances its role. Rich family background and school characteristics have a pseudo-R^2 of 18% The incremental R^2 of test scores over rich family background is a modest 5%. At the same time, the point estimate of test scores is still large: a high-scoring student is twice as likely to start college as a mid-scoring student. Conditional logit (within-school) estimates are similar.

When we use the method that permits a head-to-head competition, family background remains more important than test scores. Specifically, a one standard deviation increase in family background increases the odds of starting college by 17.8 percentage points, while a one standard deviation decrease in test scores raises the odds of dropping out by 12.8 percentage points.

Teen pregnancy: As with the educational outcomes, test scores have a large statistically significant effect on predicting which teenage women have children out of wedlock. The pseudo-R^2 is 9%, and a woman with 1 standard deviation above-average test scores is only 40 percent as likely as the mean scoring woman to have a child out of wedlock. The estimated coefficient on test scores is similar to that of H&M's SES index. At the same time, if the SES index causally precedes test scores, then SES is a more important contributor to pseudo-R^2: its R^2 alone is 7%, and the incremental R^2 of test scores is 4%.

Once again, including a richer measure of family background than did H&M greatly enhances its role. Rich family background and school characteristics have a pseudo-R^2 of 16% The incremental R^2 of test scores over race and rich family background is a modest 2%. At the same time, the point estimate of test scores is still large: a high-scoring women is only 57 percent as likely to have a child out of wedlock as a mid-scoring woman. Once again, conditional logit (within-school) estimates are similar, as are results with a head-to-head competition (comparing dP/dX's).

Arrests: Test scores also have a large statistically significant effect on predicting which young men report having been arrested. The pseudo-R^2 is 7%, and a youth with 1 standard deviation above-average test scores is half as likely as the mean scoring respondent to report having been arrested. The effects of test scores is smaller than that of Herrnstein and Murray's socioeconomic status index when the two are entered simultaneously. If the SES index causally precedes test scores, then SES is almost as important a contributor to pseudo-R^2: its R^2 alone is 4.7%, and the incremental R^2 of test scores is 5%.[7]

As usual, a richer measure of family background greatly enhances its role. Rich family background and school characteristics have a pseudo-R^2 of 42%. The incremental R^2 of test scores over rich family background is a minuscule 0.2%. At the same time, the point estimate of

[7] In contrast to the reduced form estimates we and H&M present, the structural model estimated by Jeff Grogger using the NLSY finds that test scores have no direct effect on arrest rates. He finds that high wages reduce crime, and that the capabilities measured by test scores affect crime solely through their small estimated effect on wages. He estimates that a one standard deviation increase in test scores raises wages by about 1 percent, which would lowers predicted crime per year by about .6 to .9 percent. These results would be larger if we factored in the effect of test scores on education, given that education also raises wages.

test scores is still large (though much smaller than with test scores entered along); a high-scoring student is three fourths likely to report being arrested as a mid-scoring student. Results are similar with the within-school (conditional logit) and head-to-head (dP/dX) estimates.

In results not shown, these effects were almost identical when the outcome was self-reports of whether the young man had ever been in juvenile hall. Because this self-report is plausibly more accurate, it suggests our results are not dominated by self-report biases. Results were also similar for self-reported membership in a gang: test scores had a statistically significant effect, but much lower contribution to R^2, than that of rich family background. Finally, these results follow Herrnstein and Murray in analyzing criminal activity only of young men. In analyses not shown, results predicting self-reported arrests and gang membership of young women were quite similar.

Overview of results: In predicting these varied outcomes, test scores alone almost always have a higher contribution to pseudo-R^2 than Herrnstein and Murray's socioeconomic status index does alone. Moreover, in a direct horse race, test scores have a similar standardized coefficient when compared with the H&M's SES index. These results replicate Herrnstein and Murray's basic findings that test scores are more important predictor variables than the limited elements of family background measured by the H&M SES index.

The results are more balanced when we consider that family background is largely causally prior to test scores. Typically, H&M's SES index entered alone has a pseudo-R^2 of the same magnitude as the increment to pseudo-R^2 of test scores entered after the SES index.

H&M's thesis that test scores are more important than family background is more dramatically reversed when the single SES index is replaced with a richer measure of family background. For all but the college-related outcomes, rich family background has a larger incremental contribution to R^2 when entered *after* test scores than test scores do alone. This is the extreme lower bound on the role of family background, and shows the results hold even if all of the correlation between test scores and family background is due to common parental genetic influences.

If we assume that family background is causally prior to test scores, the results are even more dramatic. Test scores provide between 0.5 and 5% increment to the pseudo-R^2; usually statistically significant, but never large. These results provide clear evidence that family background, broadly conceived, is more important (in Herrnstein and Murray's sense) in predicting outcomes than are test scores. These results replicate a number of analyses using the NLSY; these other analyses also find H&M's results are not robust to including broader measures of family background, and to relaxing the arbitrary assumptions underlying their socioeconomic status index (Fischer, et al., 1996; Dickens, et al., 1998; Korenman and Winship, 1996). At the same time, the estimated coefficient on test scores when entered simultaneously with family background often remains large, reminding us of the difficulty in evaluating the importance of test scores vs. family background when the two are correlated.

A perhaps surprising result is the relative unimportance of controlling for the students' junior high school. In general, the conditional logit estimates that create within-school estimates differ little from those using logits. The logit and conditional logit point estimates and explanatory power of the included variables usually do not differ substantially.

These results are not because schools do not differ substantially on most of the outcomes. To the contrary, the standard deviation of high schools' dropout rates is 13%, twice the mean, while for college attendance the standard deviation is 24%, two fifths the mean. (These figures are based on the NELS sample. Actual standard deviations are somewhat smaller, due to measurement error when using roughly 10-25 students per school. Nevertheless, these figures are approximately correct.)

Robustness checks

We performed a number of analyses to check the robustness of these results.

Measurement error on test scores: Classical measurement error will bias down the coefficient and explanatory power of test scores. Fortunately, as noted above, the math and reading subscores have fairly high reliabilities (.80 and .89 respectively), implying that their sum has a reliability over .90.

Augmenting the explanatory power of test scores by the standard formula for classical measurement error (that is, normally distributed and uncorrelated with the other independent variables) does not reverse the basic findings that rich family background remained a far more important predictor of outcomes than test scores. (Conversely, correcting for measurement error on family background would further strengthen the results in the text.)

Relaxing the constraints on test scores: Herrnstein and Murray's analysis rely on the long-standing (but controversial) theory that an individual's general cognitive ability is well represented by an average of their test scores. Thus, they use an unweighted average of the AFQT subtest scores (p. 594).

In fact, additional analyses of the NLSY find that a simple linear combination of test scores does not usually capture all of the useful information in the tests for predicting outcomes (Cawley, et al., 1997). These results suggest that further analysis of test scores may increase their explanatory power. In fact, in results not shown, the explanatory power of the math and reading test scores was increased when the two scores were entered separately. When the reading and math tests were entered separately, the math test typically was more important in predicting outcomes. Because no theory suggests the relationship is linear, we also entered squared terms, which further increased the explanatory power of the regressions. In separate analyses, we analyzed the sum of the four tests, including science and history along with math and reading. None of these changes reversed the basic findings that rich family background remained a far more important predictor of outcomes than test scores.

Degrees of freedom: The reliance on relative pseudo-R^2s and increment to R^2s ignores the fact that the regressions include far more explanatory variables measuring rich family background than measuring test scores. Mere chance implies a higher R^2 will follow from adding more explanatory variables, even if they have no role to play in explaining outcomes. Fortunately, the very large sample relative to the number of explanatory variables suggests that this problem is not serious.

It is possible that the broad-ranging list of variables we included to capture family background measures factors that are not properly captured by the concept of a more or less advantaged "family background." We reran results using a shorter list of variables that was closer to the concept of "advantage" -- specifically, we dropped the indicators of region, rural vs. urban, religion, gender, and birth order. (These measures were chosen on *a priori* grounds, not while examining the estimated effect sizes to pick small effects.)

While region is important for predicting idleness and religion is very important for predicting whether a youth drops out of high school, these changes do not have a large effect on the estimates of the pseudo-R^2 of the bundle of rich family background. For example, the pseudo-

R^2 of the narrower family background list plus test scores is 29.5%, which is not too far below the pseudo-R^2 of the complete list, 33.5%. The coefficient on test scores also rose by less than one standard error. For both specifications, the explained variance of the family background vector dwarfs the incremental effect of test scores. For most of the other outcome measures, the change in pseudo-R^2 was under 1 percent and the omitted variables were barely or not at all statistically significant. The fact that dropping 15 of the 51 indicators of family background had so little effect reassures us that the "horse race" is not substantially misleading due to over-inclusion of variables.

Additional outcomes: We examined several additional outcome variables in addition to those discussed in the text, including temporary dropouts, receiving a GED vs. high school diploma, and attending a two- vs. four-year college. The results were consistent with those discussed in the text.

Interactions: No theory implies that family background and tests scores operate independently. In fact, the interaction between the two is large and statistically significant negative some specifications. For example, consider the regression with rich family background and test scores. For a student with mean value on H&M's SES index, a one standard deviation increase in test scores increases the odds of attending college by 138 percent. In contrast, the same increase in test scores for a student with an SES index two standard deviations above average would enjoy only an 76 percent increase in the odds of attending college. This result may merely be capturing non-linearities in the relationship, because in the corresponding probit regression the interaction was not significant. Regardless of specification, including the interaction never made test scores more important than family background at any level of family background in the sample.

Additional functional forms: We re-estimated the regressions using probit instead of logit. We also estimated OLS and OLS with fixed school effects instead of logit and conditional logit. Results were qualitatively similar. The logit and probit pseudo-R^2's for uncommon events such as dropouts were substantially larger than the OLS R^2. The R^2 and pseudo-R^2 were much closer for common events such as starting college. In contrast, OLS coefficients were larger than the corresponding changes in probabilities estimated with probit.

Additional samples and specifications: The most common measure of general cognitive ability, IQ, is a test score adjusted for age. To perform a similar adjustment, we regressed test scores on age and used the residuals in our analyses. No results changed by a meaningful amount.

In addition, some students skip grades, while others are held back. It is possible the causal links between test scores and outcomes for these students to differ from those of typical students. Thus, we re-ran the analyses dropping students who were old or young for their grade level. No results changed by a meaningful amount.

Finally, many private schools select the students they enroll. Moreover, parents have a choice whether to send their children to private school. In both cases, it is possible the causal links between test scores and outcomes for private school students to differ from those of typical students. While not a complete analysis, we performed the simple robustness check of dropping all students who attended private schools in the eighth grade. Again, no results changed by a meaningful amount.

Decomposing Racial Differences in Outcomes

To better understand the differences in outcomes for blacks and non-Hispanic whites, we use Oaxaca decompositions. These decompositions simulate how much of the mean differences we observe between blacks and whites can be accounted for statistically by different observable characteristics of the two groups, and how much is due to different returns to those characteristics. Specifically, we estimate models with a full set of interactions between race and all of the family background and test score variables (a procedure with results identical to estimating separate regressions for each group).

We then take the coefficients estimated in the sample of black students and multiply them by the mean characteristics (family background, school characteristics, and test scores) of white students. This calculation measures the mean outcome we would predict if blacks had the same returns to characteristics as they do now, but had white family background and test scores. We then repeat this calculation with average white family background and school characteristics, and average black test scores. Finally, we repeat this calculation with average black family background and school characteristics, and average white test scores.

Permanent drop outs: In this sample, 5.2 percent of whites and 7.7 percent of blacks permanently dropped out of high school. In our simulations, if blacks had family backgrounds and school characteristics similar to whites (while keeping their black test scores), this raw difference of 2.5 percentage points higher dropout rate for blacks would reverse to a 0.85 percentage point lower dropout rate for blacks. This result is consistent with the hypothesis that disadvantaged background is responsible for all of the higher black dropout rate.

Conversely, in simulations where blacks kept their mean background but had white mean test scores, blacks are .81 percentage points more likely to drop out. With both mean white test scores and mean white backgrounds, the simulated black advantage becomes even larger: with 2.6 percentage point lower dropout rate--that is, the simulated black dropout rate would be only one half of the actual white rate (2.6% in the simulation, vs. 5.2% for whites).

Blacks who were not permanent dropouts were more likely to receive a GED than were whites (6.4% for blacks vs. 4.4 % for whites). Differences in both test scores and family background account for part of the gap. In the simulation, a black with mean white test scores and background would be more likely to receive a diploma, and less likely to receive a GED, than a similar white.

College: While 79 percent of white high school graduates had attended college by age 20, only 73 percent of blacks had. In these simulations, either family background or test scores can explain this gap. If blacks had family backgrounds and school characteristics similar to whites (while keeping their black test scores), this raw difference if 6 percentage points would shrink to approximately zero. Conversely, if blacks had test scores similar to whites, but kept their background characteristics, the difference would also shrink to approximately zero. Finally, controlling for detailed family background and school characteristics as well test scores, the black disadvantage is reversed. In fact, our results predict that a black with mean white family background and test scores has 8.4 percentage point higher probability of attending college.

A different set of findings appears if we analyze the odds of attending a four-year college instead of a 2-year college, conditional on attending college. Here, blacks who attended college were slightly more likely to attend a four-year college (64% vs. 61%). This advantage roughly doubles adjusting for differences in test scores and family background -- that is, the decompositions strengthen a mystery, but do not help solve it.

Child out of wedlock: Almost nine percent of the white women and 29 percent of the black women had a child out of wedlock by age 20. When we simulate the fertility rate for blacks if they had mean white backgrounds (and black test scores) about a third of the 20 percentage point gap goes away (the gap falls to 12 points). If we simulate the fertility rate for blacks if they had mean white test scores, but mean black backgrounds, about a fourth of the 20 percentage point gap goes away (the gap falls to 14 points). For blacks with mean white backgrounds and test scores, over half the gap is eliminated (the gap falls to 9 points).

Arrests: Almost 16 percent of the white men in this sample report having been arrested, compared to almost 22 percent of the black men. Racial differences in self-reported arrest rates are presumably due in part to different treatment of blacks and whites by the criminal justice system, and may also be influenced by different patterns of self-reporting. We are unable to measure these two potentially important factors. (Racial differences in self-reported arrest rates are discussed in Grogger, 1997.)

When we simulate black men with white family backgrounds and black test scores, slightly more than half this 6 percentage point gap goes away (the gap shrinks to 3 percentage points). In the converse case, when we simulate black men with black family backgrounds and white test scores, half this 6 percentage point gap goes away (the gap again shrinks to 3 percentage points). Finally, when we simulate black men with mean white family backgrounds and test scores, the estimated arrest rate for blacks becomes the same as that of whites. In predicting self-reported time in juvenile hall, the gap in backgrounds is somewhat more important than the gap in test scores in explaining the gap in arrest rates, but the results are roughly similar.

Racial differences in coefficients: In examining the coefficients for the various regressions, Blacks and whites usually had nearly identical returns to test scores. In contrast, family background typically mattered somewhat more for whites than for blacks. (The exception was in predicting out-of-wedlock fertility. Here, family background mattered more for blacks than for whites.) Nevertheless, the differences did not tend to be large -- that is, the simulated racial gap for a typical youth from a quite disadvantaged family was typically similar to the gap for a youth with typical white characteristics.

Robustness tests: Because the coefficients differ for the Black and white samples, our results could differ depending on the values where we perform the decomposition. To test for this sensitivity, we reran the decompositions for a disadvantaged white and disadvantaged Black, instead of for the mean white and Black. To do this, we constructed the profile of a "disadvantaged" white and disadvantaged black, moving most variables down from the mean by roughly one standard deviation.[8] We then reran the decompositions. Results were little different from those presented in the table.

We also performed the reverse decompositions: simulating white outcomes if they had black family backgrounds and /or test scores (Neumark, 1988). The results were generally consistent with those presented above, with family backgrounds and test scores each independently explaining most of the gaps in education and self-reported arrest rates, and the two together explaining a substantial portion of the fertility gap.

To check whether our linear functional form affected the results, we performed logistic and probit regressions with dummies for race. Results were similar: Simulations using all specifications had similar amounts of "over-explanation" of the estimated racial gap in educational outcomes.

Overview of the decompositions: In summary, these decompositions are consistent with the hypothesis that differences in test scores and differences in family background explain (in a statistical sense) roughly equal shares of the "explainable" black-white gaps in outcomes. Confusing any simple summary, the "explainable" share of the gap does not remain bounded by zero and 100%, but ranges from less than -200% to over +200%.

The decompositions of the racial gaps in education support more theories than they reject. The result that family background matters substantially is consistent with the hypothesis that disadvantaged background is responsible for all of Blacks' higher dropout rate, and almost all of the lower Black college attendance. This result is consistent with the U.S. not providing equality of opportunity. At the same time, it also provides no evidence for the effects of racial discrimination against youth beyond those that lead to disadvantages in family background.

Similarly, test scores "explain" statistically the black-white gaps in educational attainment. This result is consistent with the hypothesis that whatever leads to lower school and cognitive

[8] Our benchmark "disadvantaged" white had income, education, occupational status, and test scores one standard deviation below the mean. The disadvantaged youth also was male, came from a single-parent home in which the mother was unemployed, and had three siblings, of which at least one older sibling has dropped out of high school.

achievement by eighth grade also explains lower college attendance. It is, thus, consistent with notions of meritocracy, while downplaying discrimination.

The two results are both possible because background factors and test scores collectively "over-explain" the gaps in educational attainment. The result that blacks with mean white background and families are predicted to be *more* likely to finish start college provides no evidence for theories that blacks lack drive or ambition. For example, Mason interprets similar results in the NLSY as evidence of higher effort among blacks (1996). Unmeasured skills are conceptually similar to test bias against blacks -- a black youth with the same test scores as a white will, on average, have higher total ability. At the same time, this result is consistent with theories of reverse discrimination (H&M, ch. 19).

Two facts make it likely that reverse discrimination is not the main force at work. First, virtually all students get into one of their preferred colleges, even if they do not get into their first choice among highly selective college. Thus, reverse discrimination is much more likely to come into play at elite colleges than at college attendance in general (Kane, 1995). Moreover, in this dataset the "over-explanation" of the education gap is roughly the same size for both high school dropouts and college attendance. Presumably, affirmative action does not lead to reverse discrimination in high schools. (However, it is possible the over-explanation for high school graduation is due to blacks on average attending less demanding high schools.)

Similar to the education results, differences in tests scores and family background both explain (in the statistical sense) part of the gap in childbearing out of wedlock. Unlike the education results, tests scores and family background together explain only about half the racial difference in childbearing out of wedlock. These results are consistent with different cultural rewards and sanctions for fertility out of wedlock in black families and culture.

For self-reported arrests and time in juvenile hall, the results are consistent: both family background and test scores mattering, and together explain the gap.

Discussion

Our results provide no support for the hypothesis that the cognitive abilities measured by test scores are more important than family background in predicting youth's outcomes. H&M's

---

The child's mother was a teenage parent and was foreign born. The family lived in the rural, central United States, concluded that test scores predict most outcomes far better than family background. A number of replications examine the same dataset and, by relaxing H&M's arbitrary assumptions, come to the opposite conclusion. Our results strongly support those of the other replications -- family background, not test scores, are have the greatest influence on the outcomes we study. Multicollinearity, lack of precision concerning how much family background causes test scores, and other empirical problems preclude a precise analysis of the relative importance of the two sets of factors. Nevertheless, heuristically, family background is typically 4 to 10 times more important than the abilities measured by standard achievement tests in predicting youth's outcomes.

The relative unimportance of the within-school estimators suggests that family background and test scores do not primarily matter by affecting the school and neighborhood in which a youth lives. In contrast, significant past research finds important neighborhood effects on youth’s outcomes (Crane, 1991). (On the challenge of identifying context effects see Charles Manski (1993) and Robert Hauser (1970).) It is possible that our results are due to more careful measurement of family background, and that estimated neighborhood effects actually proxy for family characteristics. Nevertheless, further research is called for to reconcile these different findings. One partial explanation for the differences may be the relatively limited between-school variance in mean achievement; it could be that schools have larger effects on the speed of learning, where the between-school variance is higher (Bryk and. Raudenbush, 1988.)

Turning to racial differences, more than half the racial differences in education and in self-reported arrests can be explained (statistically) by differences in family background. More than half the racial differences can also be explained (statistically) by differences in test scores.

In this study we do not name the racial gaps we identify; such gaps are a measure of our ignorance. Correlational results, particularly concerning difficult-to-measure constructs such as ability and discrimination, are rarely easy to interpret. In fact, the results presented here (as are those in *The Bell Curve*) are consistent with discrimination (or reverse discrimination), biased tests, or unmeasured skills, culture, and energy among Black youth.

An implication of the common "over-explanation" result is that many analysts follow a procedure likely to lead to false acceptance of their theory. Because decompositions such as

---

and the child attended a school which the teacher claims is dangerous and has a problem with drugs on campus.

those pursued in this paper can "explain" over 100% of a phenomenon, factors that explain 75% should not be considered "clearly the most important factor. Consistent with some past research, but in contrast to the interpretation of H&M, factors that “explain” 100% of the racial gap in an outcome can explain far less than 100% of the *explainable* gap (given that test scores and family background together explain over 100% of the racial education gap). A more complete understanding is required to make such statements, and this understanding remains lacking.

The naming of the residual black-white wage gap in a wage equation as "discrimination" is known to be problematic. Omitted variables that are correlated with race and that affect productivity or working conditions may cause the wage gap. Thus, a researcher who finds an additional variable that diminishes the pay gap to zero may claim evidence that discrimination does not exist. Unfortunately, the equation still may omit many important variables that are correlated with race and that affect productivity. If such additional variables were measured, measured discrimination might reappear.

Further research should untangle which aspects of family background are most important for each outcome. Recalling that many of these measures are themselves proxies for more complicated behaviors within the family, these results must be supplemented with additional qualitative research on what makes children succeed.

In the 1995 “Afterword” to *The Bell Curve*, Murray emphasizes what he feels is the bottom line of “perhaps the most important section” (p. 567) the book: “The kinds of economic and social disadvantages that have been treated as decisive in recent discussions of social policy are comparatively unimportant....[This result undermines] the rationale for many of the social policies that came into vogue in the 1960s.” (p. 572) As noted above, H&M’s measure of “comparatively unimportant” do not use the cost-benefit framework advocated by most policy analysts; thus, their results do not shed much light on evaluating the rationale for social policies. At the same time, H&M pose a puzzle worth investigating. If family and neighborhood background have little influence on youths’ outcomes, then policies that make up for deficiencies in such backgrounds have little scope to be effective.

The original contribution of this paper is to analyze a dataset different from the one H&M and previous replications have examined. We support several previous replications that find that, in contrast to Murray’s claim, the kinds of economic and social disadvantages that have been treated as decisive in recent discussions of social policy are comparatively *important*. Identifying cost-effective policies remains a challenge; nevertheless, our results strongly suggest that policies that can make up for differences in family and neighborhood background have scope to be effective.

Table 1: Summary Statistics			
	Entire Sample (N = 14662)	Whites (N = 10073)	Blacks (N = 1496)
	Mean s.d.	Mean s.d.	Mean s.d.
Permanent dropout (Omits those with GED and those who started college with no HS diploma)	0.061	0.052	0.077
Attended college (Among HS graduates)	0.785	0.791	0.729
Obtained a GED instead of diploma (Among those with either)	0.047	0.044	0.064
Had a child out of wedlock (Among women)	0.120	0.087	0.289
Has been arrested (Among men)	0.169	0.158	0.217
Is currently not working and not in school (Among men)	0.112	0.096	0.222
Attended four-year college, not two-year (Among those with some college)	0.594	0.607	0.637
(Variables below here --other than H&M's SES Index and measures including test scores -- are included as "rich family background" in Tables 2 and 3.)			
Income (z)	-0.036 (1.036)	0.135 (0.917)	-0.719 (1.184)
Father's education (z)	0.005 (0.899)	0.088 (0.871)	-0.143 (0.646)
Mother's education (z)	0.006 (0.959)	0.102 (0.884)	-0.087 (0.850)
Father's occupational status (z)	0.003 (0.872)	0.059 (0.885)	-0.182 (0.722)
Mother's occupational status (z)	0.004 (0.917)	0.077 (0.913)	-0.200 (0.948)
H&M's SES Index (z)	0.015 (1.038)	0.175 (0.960)	-0.426 (0.936)
Cognitive Test Scores (z)	0.004 (1.000)	0.165 (0.974)	-0.605 (0.826)
Interaction between test scores and SES	1.101 (2.513)	1.038 (2.446)	1.197 (2.404)
Squared z-score of income	1.075 (2.321)	0.859 (1.789)	1.919 (3.409)
Number of siblings	2.268 (1.574)	2.099 (1.474)	2.696 (1.807)
Missing father's education	0.188	0.154	0.420
Missing mother's education	0.073	0.060	0.075
At least one parent went to college	0.667	0.707	0.596
At least one parent dropped out of high school	0.284	0.225	0.312
Father unemployed	0.063	0.050	0.090
Mother unemployed	0.275	0.268	0.261
Oldest child	0.299	0.312	0.285
Male	0.478	0.484	0.454
Religion - Baptist (Missing religion is non-Baptist Protestant)	0.187	0.177	0.537
Catholic	0.311	0.288	0.072
Other	0.165	0.137	0.176
Jewish	0.016	0.022	0.001

No religion reported	0.036	0.032	0.033
No religion	0.027	0.026	0.015
Very religious (Missing category is not religious)	0.397	0.409	0.394
Religious	0.150	0.154	0.166
Somewhat religious	0.161	0.165	0.147
Step-parent in the home (Missing category is two-parent family)	0.194	0.182	0.232
Single parent family	0.159	0.136	0.396
A parent remarried between 8th and 10th grade	0.079	0.077	0.118
A parent remarried between 10th and 12th grade	0.050	0.053	0.071
Native speaker of English	0.796	0.889	0.831
Father foreign born	0.210	0.094	0.137
Mother foreign born	0.210	0.091	0.153
Parents worked at least a combined 40 hours while in 8th grade	0.264	0.211	0.467
Parents worked at least a combined 40 hours while in 12th grade	0.492	0.537	0.372
An older sibling dropped out of high school	0.084	0.077	0.102
At least one parent was a teenage parent	0.252	0.248	0.301
More than 50 books in home	0.879	0.918	0.782
Has at least one magazine subscription	0.744	0.808	0.623
Family has a public library card	0.757	0.780	0.701
Lived in South (Missing region is Northeast)	0.351	0.321	0.636
Lived in West	0.196	0.144	0.064
Lived in central city	0.269	0.328	0.149
Lived in an urban area (Missing category is suburban)	0.254	0.195	0.379
Lived in rural area	0.313	0.353	0.308
School teaching is considered good (Student report)	0.801	0.801	0.783
School is dangerous (Teacher report)	0.534	0.465	0.684
Drugs are a problem in school (Teacher report)	0.443	0.420	0.390
Parents involved in school organizations like PTA	0.509	0.530	0.517
The youth was a member of some club like boy scouts when younger	0.819	0.887	0.760
There are rules about the amount of TV watched	0.830	0.857	0.817
There are rules about household chores and homework	0.890	0.906	0.880
Parents help with homework	0.387	0.403	0.429
School has less than 25% single parent families	0.447	0.485	0.301
School has between 25 and 50% single parent families	0.365	0.361	0.402
School is less than 40% minority	0.758	0.912	0.371

Notes: Variables marked (z) are z-coded to have mean zero and standard deviation equal to one. Construction of H&M's SES Index (Herrnstein and Murray's index of socioeconomic status) is described in the text.

Table 2: The Effects of Test Scores and Family Background on Outcomes
Logistic regressions, Odds-Ratios and pseudo-R^2's

	High School Dropout (no GED)	Started College by Age 20	Obtained a GED instead of a Diploma	Had a Child out of wedlock
Sample (White)	Omits those with GED and those who started college with no HS diploma	High school graduates	All who obtained either a GED or HS diploma	Women
1. Test scores				
Odds-ratio, test scores	.237	2.92	.534	.401
Pseudo-R^2	.157	.133	.038	.085
2. H&M SES index				
Odds-ratio, SES index	.301	3.03	.536	.458
Pseudo-R^2	.136	.129	.037	.068
3. H&M SES index & test scores				
Odds-ratio, test scores	.325	2.34	.628	.496
Odds-ratio, SES index	.431	2.38	.637	.605
Pseudo-R^2	.211	.197	.054	.107
Inc. R^2 of test scores [Rows (3) -(2)]	.075	.068	.017	.039
4. Rich family background				
Pseudo-R^2	.295	.185	.379	.162
5. Rich family background & test				
Odds-ratio, test scores	.403	2.20	.807	.572
dP/dX of test scores	-.044	.128	-.009	-.044
dP/dX of family background	-.068	.178	-.058	-.071
Pseudo-R^2	.335	.235	.381	.183
Inc. R^2 of test scores [Rows (5) -(4)]	.040	.050	.002	.021
Inc. R^2 of family bkgnd [Rows (5) -(1)]	.178	.102	.343	.098

Notes: Odds Ratios for Test Scores and for SES measure the change in the odds of an outcome when the test scores or H&M SES index change one standard deviation. An odds ratio of 0.9 for Test scores, for example, indicates the outcome is 90% as likely for high-scoring students--for example, 20% of the baseline respondents, but 18% for students with high test scores

Incremental R^2 of test scores is the pseudo-R^2 of test scores and family background or SES index minus the pseudo-R^2 of family background or SES index alone. Incremental R^2 of family background or SES index are the R^2 of test scores and either family background or SES index minus the R^2 of test scores alone.

H&M SES Index is a single measure averaging parents' occupational status, parents' education, and family income. Rich family background includes a number of family and school characteristics, listed in Table 1. dP/dX estimates the effect on the probability of the outcome of a one standard deviation increase in test scores or in the bundle of family characteristics.

Table 2: The Effects of Test Scores and Family Background on Outcomes
Logistic regressions, Odds-Ratios and pseudo-R^2's

	Had been arrested	Is currently not working and not attending college (age 20)	Attended a four year instead of a two year college
Sample (White)	Men	Men	Attended some College
1. Test scores			
Odds-ratio, test scores	.469	.630	2.15
Pseudo-R^2	.073	.027	.080
2. H&M SES index			
Odds-ratio, SES index	.546	.543	1.88
Pseudo-R^2	.047	.044	.052
3. H&M SES index & test scores			
Odds-ratio, test scores	.539	.767	1.90
Odds-ratio, SES index	.690	.605	1.56
Pseudo-R^2	.087	.051	.101
Inc. R^2 of test scores [Rows (3) -(2)]	.050	.007	.049
4. Rich family background			
Pseudo-R^2	.424	.091	.097
5. Rich family background & test			
Odds-ratio, test scores	.797	.808	1.83
dP/dX for test scores	-.030	-.018	.144
dP/dX for family background	-.223	-.065	.157
Pseudo-R^2	.427	.094	.135
Inc. R^2 of test scores [Rows (5) -(4)]	.003	.003	.038
Inc. R^2 of family bkgnd [Rows (5) -(1)]	.354	.067	.055

Table 3: Oaxaca Decompositions of Racial Differences in Outcomes
Outcome measures are presented as percentages.

	High School Dropout*	Started College by age 20*	Obtained GED instead of Diploma*	Childbirth out of wedlock
1. White rate	5.20	79.07	4.39	8.68
2. Black rate	7.77	72.93	6.43	28.91
Gap: Black - white [row (2) - (1)]	2.57	-6.14	2.04	20.23
3. Simulated Black rate with white test scores	4.39	78.62	6.29	22.94
Gap: Black simulation - white actual [row (3) - (1)]	-0.81	-0.45	1.90	14.26
4. Simulated Black rate with white family background	6.05	79.94	2.60	20.23
Gap: Black simulation - white actual [row (4) - (1)]	0.85	0.87	-1.79	11.55
5. Simulated Black rate with white test scores and family background	2.57	87.45	2.43	17.18
Gap: Black simulation - white actual [row (5) - (1)]	-2.63	8.38	-1.96	8.50
Percentage of Black-White Gap Accounted For	202.33	236.48	196.08	57.98

Notes: All "gaps" are simulations from the Oaxaca decompositions described in the text. Figures represent the predicted percentage point gap if Blacks had the same returns to the characteristics listed in each row as do Whites. The sample for each outcome is described in Table 2.

* Outcomes marked with a star exhibit "over-explanation" -- the sign of the simulated black-white gap is the reverse of the sign of the raw gap.

Table 3: Oaxaca Decompositions of Racial Differences in Outcomes (continued)

Outcome measures are presented as percentages.

	Had been arrested*	Is currently not working or in school (age 20)	Attended a four year instead of a two year college
1. White rate	15.76	9.57	60.66
2. Black rate	21.79	22.24	63.67
Gap: Black - white [row (2) - (1)]	6.03	12.67	3.01
3. Simulated Black rate with white test scores	18.67	20.44	68.26
Gap: Black simulation - white actual [row (3) - (1)]	2.91	10.87	7.60
4. Simulated Black rate with white family background	18.39	14.39	62.02
Gap: Black simulation - white actual [row (4) - (1)]	2.63	4.82	1.36
5. Simulated Black rate with white test scores and family background	15.72	12.86	70.87
Gap: Black simulation - white actual [row (5) - (1)]	-0.04	3.29	10.21
Percentage of Black-White Gap Accounted For	100.66	74.03	-239.20

References

Bryk, Anthony S., and Stephan W. Raudenbush. 1988. "Toward a More Appropriate Conceptualization of Research on School Effects: A Three-Level Hierarchical Linear Model." American Journal of Education, 97:65-108.

Cawley, John, Karen Conneely, James Heckman, and Edward Vytlacil, "Measuring the Effects of Cognitive Ability," this issue.

Chatman, Jennifer A. and Charles A. O'Reilly, III, "Worker Smarter and Harder: A Longitudinal Study of Managerial Success," Administrative Science Quarterly, v. 39, 1994, pp. 603-627.

Crane, Jonathan "Effect of Neighborhoods on Dropping out of School and Teenage Childbearing," in Jencks, Christopher and Paul E. Peterson, eds., The Urban Underclass, Brookings, DC 1991.

Currie, Janet and Duncan Thomas, "Race, Children's Cognitive Achievement and *The Bell Curve*," this issue.

Dickens, William T., Thomas Kane and Charlie Schultz "What do We Know About the Extent of Race Discrimination?" This issue.

Farley, Reynolds, and Frey, William H. 1992. "Changes in the Segregation of Whites from Blacks During the 1980s: Small Steps Towards a More Racially Integrated Society," Research Report, 92-257, Population Studies Center, University of Michigan.

Fischer, Claude, Michael Hout, and others, Inequality by Design, Princeton University, Princeton NJ, 1997.

Gamoran, Adam., "Measuring curriculum differentiation," American Journal of Education, 1989 Feb., v. 97, n. 2:129-143.

Grogger, Jeff, "Market Wages and Youth Crime," NBER Working Paper 5983, Cambridge MA, March 1997.

Hauser, Robert M. 1970. "Context and Context: A Cautionary Tale." American Journal of Sociology, 75:645-664.

Haveman, R. and B. Wolfe, Succeeding Generations: On the Effects of Investments in Children, NY Russell Sage, 1994.

Herrnstein, Richard and Charles Murray, The Bell Curve, Free Press, NY, 1995. [H&M]

Hout, Michael, Claude Fischer, and Richard Arum, "IQ, Poverty, and *The Bell Curve* Controversy, " this issue.

Hsiao, Cheng, Analysis of Panel Data, Cambridge U. Press, Cambridge, UK, 1986.

Juhn, C., K. Murphy, B. Pierce, "Accounting for the slowdown in Black-White Wage Convergence," in M. Kosters, ed., Workers and their Wages, Washington DC, AEI Press, 1991.

Kane, Thomas, "Racial Preferences in College Admissions," mimeo, Brookings, 1995.

Kingsley, G. Thomas and Margery Austin Turner, editors. Housing Markets and Residential Mobility, Washington, D.C. Urban Institute Press, Lanham, MD: 1993.

Korenman, Sanders, and Steven Winship, "A Re-analysis of *The Bell Curve*," NBER w.p. 5230, Sept. 1995.

Lucas, Samuel R., and Adam Gamoran, "Race and Track Assignment," mimeo, U. of Wisconsin-Madison, 1991.

Manski, Charles F. 1993. "Identification Problems in the Social Sciences." Sociological Methodology, 23:1-56.

Mason, Patrick L., "Race Culture and Skill: Interracial Wage Differences Among African Americans, Latinos, and Whites," mimeo, Dept. of Economics, University of Notre Dame, 1996.

Maxwell, N., "The Effect on Black White Wage Differentials of Differences in the Quality and Quantity of Education, Industrial and Labor Relations Review, 47, 2, Jan 1995, pp. 249-264.

National Center for Education Statistics, Documentation of the NELS found at http://www.ed.gov/NCES/pubs/coe95/source14.html, 1995.

National Center for Education Statistics, National Education Longitudinal Study of 1988, Second Follow-up: Student Component Data File User's Manual, Washington DC, June 1994

Neumark, David, "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," Journal of Human Resources, 23, 1988, pp. 279-295.

Oaxaca R.L, and Ransom M.R., "On Discrimination and the Decomposition of Wage Differentials," Journal of Econometrics, 1994 Mar, v. 61 n. 1, pp. 5-21.

Rock, Donald A. and Judith M. Pollach, The Psychometric Report for the NELS Base Year Test Through Second Follow-Up, Washington, D.C., U.S. Dept. of Education, Office of Educational Research and Improvement, (National Center for Education Statistics report 95-382), 1995.

Appendix 1
Logistic Regression
Dependent Variable = Permanent Dropout
(N = 9663)
Sample includes whites without a GED

	Rich Family Background		Rich Family Background and Cognitive Test Scores	
	Odds Ratio	Std Dev	Odds Ratio	Std Dev
Income (z)	0.777	0.086	0.827	0.093
Squared z-score of income	0.930	0.035	0.922	0.036
Father's education (z)	0.723	0.086	0.805	0.096
Mother's education (z)	0.647	0.065	0.738	0.077
Missing father's education	0.893	0.209	0.859	0.209
Missing mother's education	1.334	0.378	1.223	0.357
At least one parent dropped out of high school	1.051	0.169	1.078	0.176
Father's occupational status (z)	0.881	0.072	0.906	0.075
Mother's occupational status (z)	0.909	0.059	0.945	0.063
Father unemployed	1.137	0.212	1.141	0.219
Mother unemployed	1.081	0.134	1.092	0.139
Number of siblings	1.165	0.161	1.240	0.174
Oldest child	0.860	0.091	0.803	0.088
Male	1.269	0.197	1.227	0.193
Religion - Baptist (Missing religion is non-Baptist Protestant)	0.862	0.143	0.840	0.142
Catholic	1.594	0.293	1.682	0.315
Other	1.476	0.365	1.408	0.361
No religion reported	1.060	0.287	1.008	0.282
No religion	0.202	0.031	0.231	0.036
Very religious (Missing category is not religious)	0.361	0.060	0.409	0.069
Religious	0.261	0.046	0.292	0.053
Somewhat religious	1.094	0.039	1.078	0.039
Step-parent in the home (Missing category is two-parent family)	1.554	0.225	1.622	0.235
Single parent family	2.066	0.569	2.186	0.618
A parent remarried between 8th and 10th grade	1.279	0.221	1.340	0.235
A parent remarried between 10th and 12th grade	1.951	0.370	1.845	0.357
Father foreign born	0.467	0.155	0.427	0.152
Mother foreign born	1.102	0.335	1.120	0.358
Parents worked at least a combined 40 hours while in	0.792	0.148	0.798	0.151

8th grade				
Parents worked at least a combined 40 hours while in 12th grade	0.723	0.086	0.758	0.090
An older sibling dropped out of high school	2.416	0.362	2.562	0.387
At least one parent was a teenage parent	0.921	0.125	0.907	0.126
More than 50 books in home	0.722	0.102	0.859	0.124
Has at least one magazine subscription	0.695	0.083	0.784	0.095
Family has a public library card	0.737	0.089	0.829	0.101
Lived in South (Missing region is Northeast)	1.202	0.206	1.131	0.195
Lived in West	0.797	0.161	0.738	0.153
Lived in central city	0.997	0.163	0.924	0.152
Lived in an urban area (Missing category is suburban)	1.192	0.185	1.363	0.220
Lived in rural area	0.821	0.104	0.841	0.108
School teaching is considered good (Student report)	0.644	0.075	0.745	0.088
School is dangerous (Teacher report)	1.066	0.126	0.970	0.118
Drugs are a problem in school (Teacher report)	1.173	0.138	1.166	0.141
Parents involved in school organizations like PTA	0.770	0.093	0.784	0.097
The child was a member of some club like boy scouts	0.692	0.104	0.662	0.103
There are rules about the amount of TV watched	1.029	0.165	0.969	0.160
There are rules about household chores and homework	1.334	0.317	1.446	0.368
Parents help with homework	1.297	0.146	1.037	0.121
School has less than 25% single parent families	1.979	0.331	1.967	0.331
School has between 25 and 50% single parent families	1.049	0.192	1.090	0.200
School is less than 40% minority	0.559	0.089	0.637	0.103
Cognitive Test Scores (z)			0.403	0.032

Appendix 2
OLS Results for Oaxaca Decompositions
Dependent Variable = Permanent Dropout
(N = 11057)
Sample includes whites and blacks without a GED

Adjusted R^2 = 0.219	Estimate	Std Dev
Income (z)	-0.008	0.003
Squared z-score of income	0.003	0.002
Father's education (z)	0.006	0.004
Mother's education (z)	-0.006	0.003
Missing father's education	-0.037	0.016
Missing mother's education	-0.023	0.019
At least one parent dropped out of high school	0.030	0.010
Father's occupational status (z)	-0.001	0.003
Mother's occupational status (z)	-0.001	0.003
Father unemployed	0.029	0.015
Mother unemployed	0.002	0.006
Number of siblings	0.012	0.005
Oldest child	-0.009	0.004
Male	0.014	0.007
Religion - Baptist (Missing religion is non-Baptist Protestant)	-0.005	0.005
Catholic	0.020	0.009
Other	0.009	0.015
No religion reported	0.006	0.018
No religion	-0.073	0.006
Very religious (Missing category is not religious)	-0.065	0.008
Religious	-0.076	0.007
Somewhat religious	0.004	0.002
Step-parent in the home (Missing category is two-parent family)	0.019	0.009
Single parent family	0.035	0.022
A parent remarried between 8th and 10th grade	0.015	0.011
A parent remarried between 10th and 12th grade	0.040	0.015
Father foreign born	-0.029	0.009
Mother foreign born	-0.006	0.011
Parents worked at least a combined 40 hours while in 8th grade	0.009	0.015
Parents worked at least a combined 40 hours while in 12th grade	-0.009	0.005
An older sibling dropped out of high school	0.098	0.015
At least one parent was a teenage parent	-0.010	0.006
More than 50 books in home	-0.029	0.012

Has at least one magazine subscription	-0.022	0.008
Family has a public library card	-0.017	0.007
Lived in South (Missing region is Northeast)	0.004	0.006
Lived in West	-0.015	0.007
Lived in central city	-0.002	0.006
Lived in an urban area (Missing category is suburban)	0.010	0.005
Lived in rural area	-0.010	0.005
School teaching is considered good (Student report)	-0.016	0.006
School is dangerous (Teacher report)	-0.002	0.005
Drugs are a problem in school (Teacher report)	0.007	0.005
Parents involved in school organizations like PTA	-0.010	0.005
The child was a member of some club like boy scouts	-0.045	0.013
There are rules about the amount of TV watched	-0.003	0.008
There are rules about household chores and homework	0.011	0.010
Parents help with homework	0.001	0.004
School has less than 25% single parent families	0.035	0.006
School has between 25 and 50% single parent families	0.001	0.006
School is less than 40% minority	-0.035	0.010
Cognitive Test Scores (z)	-0.028	0.002
Income (z) * black dummy variable	-0.013	0.016
Squared z-score of income * black dummy variable	-0.005	0.005
Father's education (z) * black dummy variable	0.014	0.017
Mother's education (z) * black dummy variable	0.014	0.010
Missing father's education * black dummy variable	0.035	0.031
Missing mother's education * black dummy variable	0.030	0.049
At least one parent dropped out of high school * black dummy variable	0.000	0.026
Father's occupational status (z) * black dummy variable	-0.006	0.011
Mother's occupational status (z) * black dummy variable	-0.001	0.009
Father unemployed * black dummy variable	-0.059	0.032
Mother unemployed * black dummy variable	0.025	0.020
Number of siblings * black dummy variable	0.003	0.017
Oldest child * black dummy variable	-0.012	0.015
Male * black dummy variable	-0.012	0.019
Religion - Baptist * black dummy variable	0.005	0.028
Catholic * black dummy variable	-0.029	0.028
Other * black dummy variable	-0.020	0.054
No religion reported * black dummy variable	-0.003	0.067
No religion * black dummy variable	0.008	0.021
Very religious * black dummy variable	0.002	0.024
Religious * black dummy variable	0.034	0.025

Somewhat religious * black dummy variable	0.005	0.005
Step-parent in the home * black dummy variable	0.022	0.025
Single parent family * black dummy variable	-0.057	0.043
A parent remarried between 8th and 10th grade * black dummy variable	-0.005	0.026
A parent remarried between 10th and 12th grade * black dummy variable	-0.046	0.032
Father foreign born * black dummy variable	0.000	0.034
Mother foreign born * black dummy variable	0.014	0.036
Parents worked at least a combined 40 hours while in 8th grade * black dummy variable	0.033	0.034
Parents worked at least a combined 40 hours while in 12th grade * black dummy variable	0.015	0.017
An older sibling dropped out of high school * black dummy variable	-0.053	0.035
At least one parent was a teenage parent * black dummy variable	0.038	0.019
More than 50 books in home * black dummy variable	0.015	0.023
Has at least one magazine subscription * black dummy variable	0.054	0.017
Family has a public library card * black dummy variable	0.019	0.022
Lived in South (Missing region is Northeast) * black dummy variable	0.011	0.022
Lived in West * black dummy variable	0.018	0.029
Lived in central city * black dummy variable	0.049	0.029
Lived in an urban area * black dummy variable)	-0.001	0.016
Lived in rural area * black dummy variable	0.031	0.020
School teaching is considered good * black dummy variable	-0.006	0.020
School is dangerous * black dummy variable	-0.011	0.017
Drugs are a problem in school * black dummy variable	0.005	0.017
Parents involved in school organizations like PTA * black dummy variable	-0.014	0.018
The child was a member of some club like boy scouts * black dummy variable	0.027	0.025
There are rules about the amount of TV watched * black dummy variable	0.013	0.026
There are rules about household chores and homework * black dummy variable	-0.019	0.034
Parents help with homework * black dummy variable	-0.001	0.016
School has less than 25% single parent families * black dummy variable	0.008	0.021
School has between 25 and 50% single parent families * black dummy variable	-0.009	0.017
School is less than 40% minority * black dummy variable	0.020	0.018
Cognitive Test Scores (z) * black dummy variable	-0.017	0.009
Black dummy variable	-0.182	0.059
Constant Term	0.209	0.025