Accounting for the Effect of Schooling and Abilities in the Analysis of Racial and Ethnic Disparities in Achievement Test Scores[1]

James J. Heckman
University of Chicago and
American Bar Foundation

Maria I. Larenas
Center for Social Program Evaluation

Sergio Urzua
University of Chicago

First Draft, February 29, 2004

This Draft, September 9, 2005

[1]We thank Pedro Carneiro, Dimitriy Masterov and Salvador Navarro for their helpful comments. Comments on this draft are welcomed. The authors' email addresses are: James J. Heckman, jjh@uchicago.edu; Maria I. Larenas, ??@??; Sergio Urzua, surzua@uchicago.edu.

Abstract

This paper examines the sources of racial and ethnic disparity in achievement test scores. We estimate the effect of schooling on achievement test scores and question the importance of genetic differences. In addition, we estimate the effect of family background and school inputs and other environments on test scores. Using NLSY data on schooling and test scores, we estimate the effect of schooling on test scores using the two methods developed by Hansen et al. (2004). These methods control for the endogeneity of schooling by postulating that both schooling and test scores are generated by a common latent ability. The identification strategy uses scores from tests taken at different schooling levels where individuals attain different levels of final schooling. Our findings show that schooling affects measured ability in a significant way. Larger causal effects of schooling on test scores are found among whites than among minorities. Although schooling raises test scores, it does not reduce test score gaps between blacks, whites and Hispanics. Initial test scores gaps persist regardless of schooling level. These results persist after controlling for family characteristics and school quality. This indicates that policies targeted solely in raising schooling levels to reduce racial-ethnic test score gaps might prove ineffective.

1 Introduction

Cognitive ability matters in social performance. It is established that reduction in the racial-ethnic gap in scholastic ability can help reduce disparities in social outcomes. Neal and Johnson (1996) document that cognitive skill differences between black and white teenagers explain a significant portion of the observed gap in wages. The sources of racial-ethnic disparities in wages are diverse, and while cognitive ability is not the only source of the observed gap in wages and income, it plays an important role in social success. However, the magnitude of the effect of cognitive ability and its causes remain a controversy. The main reason is the difficulty in disentangling the effects of factors that influence both cognitive ability and social outcomes. Herrnstein and Murray (1994) suggest that a single factor, genetic ability, is the only determinant of both cognitive ability and social performance. This view implies that policy interventions will be ineffective for reducing racial inequality. This paper challenges Herrnstein and Murray's view and focuses on the important role of schooling in increasing measured cognitive ability. In particular, education has a contrasting impact upon measured scholastic ability for different racial-ethnic groups. If measured ability is fixed at early ages and invariant with schooling and other interventions, policy targeted at later ages would be ineffective in reducing disparities. In contrast, if schooling increases measured ability, policies that increase education attainment and quality would be important for reducing social disparities.

Increasing schooling levels for minorities has been viewed as an important avenue for policy to reduce negative social outcomes. The effectiveness of this policy lies in the disparate impacts of schooling on measured ability (and other skills) across racial-ethnic groups and different socioeconomic groups. The effect of schooling on achievement-test scores has been the focus of extensive research (Ceci, 1991; Winship and Korenman, 1997; Hansen, Heckman, and Mullen, 2004). A substantial racial-ethnic gap in test scores has been extensively documented even among whites and minorities with the same schooling level at the time of the test. Several explanations, other than differences in genetic material, have been offered by previous research as causes for racial and ethnic disparities in scholastic achievement (Phillips, Brooks-Gunn, Duncan, Klebanov, and Crane, 1998; see Ferguson, 2002b for a review of previous literature). One major source of concern is the fact that racial differentials in cognitive skills develop at early ages and tend to widen over the life cycle (Ferguson, 2002b; Carneiro, Heckman, and Masterov, 2005; Phillips, Brooks-Gunn, Duncan, Klebanov, and Crane, 1998).

Some authors have been able to reduce the racial-ethnic testing gap using various covariates. Using data from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS), Fryer and Levitt (2004) found a reduction in the gap attributable to school quality among blacks and whites when they entered kindergarten. Using endogenous variables, Duncan and Brooks-Gunn (1997) present evidence that socioeconomic status and family income matters in explaining the gap. Using a large set of family background measures from Children NLSY, Phillips, Brooks-Gunn, Duncan, Klebanov, and Crane (1998) document that these factors explain about half of the black-white test score gap among five- and six-year-olds. Overall, the large body of literature does not achieve consensus on causes of the gap. Even after controlling for socioeconomic status, family factors, neighborhood environments and school quality, a substantial racial-ethnic gap in test scores remains.[1]

This paper applies two methodologies of Hansen, Heckman, and Mullen (2004) to identify the causal effect of schooling on test scores for blacks, Hispanics and whites. One of the major shortcomings of previous attempts is that they fail to control for differences in latent ability, or more precisely, differences in early environments that are not observable to the econometrician. The major problem with failing to control for latent ability differences is the endogeneity of schooling attainment. Latent ability influences schooling attainment and measured ability (i.e., test scores), and schooling, in turn, affects measured ability. This is a crucial distinction, as shown by Neal and Johnson (1996) who find that the presence of a racial-ethnic gap in cognitive ability will likely overstate the effect of current labor discrimination on black-white wage gaps. However, the reverse causality of latent ability and schooling will likely bias wage gaps estimates from reduced forms that only account for test scores.

To address the reverse causality of schooling, cognitive ability is modeled as manifestations of schooling and latent ability, in turn schooling does not influence latent ability. Using longitudinal data we measure the effects of different levels of attained schooling at the date of the test on

[1]Although the CNLSY and ECLS provide useful data to explain racial gaps in test scores, such as an initial measure of ability, maternal observed-ability measures and parental practices, we are unable to relate data from the CNLSY and ECLS to performance in the labor market and educational attainment.

achievement on the AFQT test, we are able to isolate the causal effect of schooling on test scores. Also, we identify the effect of covariates, such as school quality, socioeconomic and family factors on test scores. The methodology based in factor analysis, allow us to estimate the effects of exogenous interventions on latent ability for different racial-ethnic groups at different levels of the ability distribution. This also provides a crucial policy instrument for understanding and targeting policy to close the achievement gap in test scores.

Our estimates suggest a large causal effect of schooling on test scores. Although schooling raises test scores, it does not reduce test score gaps between blacks, whites and Hispanics. Environmental factors have been influential for a crucial part of the development of cognitive ability when the AFQT measure was taken. For all demographic groups, initial (ninth grade) test score gaps are maintained regardless of schooling level. In fact, larger effects of schooling are found among whites rather than among minorities. Even if favorable family environments are related to higher latent ability, failing to account for early disparities in environments (not only related to family) could overestimate the importance of genetic differences in explaining disparities in scholastic ability. Our results show that family background and school quality do not play an important role in explaining the growth of test scores among blacks, whites and Hispanics. This indicates that policies targeted only at schooling attainment might prove ineffective.

The paper is organized as follows. Section 2 presents the data, basic facts about the racial-ethnic achievement gap in test scores and its relation to schooling. We also present evidence on the importance of environmental factors, especially differences in environments during early childhood, on the achievement gap. In Section 3 we present a nonparametric control function model—proposed by Hansen, Heckman, and Mullen (2004)—that controls for reverse causality of education decisions and latent ability. Section 4 presents estimates of the causal effect of schooling for the model presented in Section 3. We analyze the impact of family environments and school characteristics on test scores. In Section 5 we compare these results with estimates from a semiparametric model that controls for the endogeneity of schooling decisions. This allows us to estimate the distribution of latent ability and estimate the causal effect of schooling at various levels of the ability distribution. In addition, we analyze the effect of school quality under different specifications. Finally, we present our conclusions.

2 Data and Basic Facts on the Test Scores Gap

In this section, we begin by describing the data and some facts about racial and ethnic disparities in test scores and environments. We use NLSY79 basic wave and additional samples of blacks and Hispanics[2]. Appendix A describes in detail the sample criteria and the variables used. The NLSY collects data on educational attainment, socioeconomic status, school characteristics, labor market outcomes, family background, family income and environments. Our sample includes individuals from 16 to 23 years old, and includes 4,200 whites, 2,429 blacks, and 1,494 Hispanics. NLSY includes information on the Armed Services Vocational Aptitude Battery (ASVAB), administered in 1980 to all respondents aged 16 to 23. The ASVAB consists of ten tests developed by the military to measure aptitude for the armed forces training programs. The Armed Forces Qualifying Test (AFQT) is the most commonly used combination of the tests. We identify six categories of schooling at the time of the test: 9th grade or lower, 10th grade, 11th grade, 12th grade, 13 to 14 years, and 15 years or more. We observed final education status by year 2000, by which time most of the sample had finished their education. Final schooling level is divided into four levels: high-school dropouts and GEDs, high-school graduates (with no postsecondary education), some college and associate degrees (two-year college degrees) and college completed or more (four or more years of college). GED recipients are considered to have completed 12 years of schooling at the test date and are grouped with high-school dropouts. Sample sizes for each level of schooling at test date and final schooling are presented in Appendix A for each demographic group (Appendix A, Table A1).

At the time of the test, 37 percent of the sample had finished high school, 20 percent had some form of postsecondary education and 43 percent had completed less than high school. Table 1 presents education levels at the time of the test and final schooling status. In 1980, at the time of test administration, some individuals were still enrolled in school and some had finished school or were in transition. Racial and ethnic gaps in educational attainment at the time of the test are evident in Panel A. College attainment is more predominant among whites, and whites attended college at a higher rate than minorities (Panel C). As shown in Panel D, the gap in test scores is

[2]The NLSY disadvantaged white (nonblack non-Hispanic) sample is excluded.

substantial for each test in the ASVAB components of the AFQT between whites and minorities.

The disparities in mean test scores do not result from comparing high-ability whites with low-ability minorities. The gap permeates the complete distribution of test scores. Figure 1 shows the distributions of age-adjusted AFQT scores for blacks, whites and Hispanics who were enrolled in school at the test date. In this figure, disparities in scholastic ability are clear. While 30 percent of whites fall below the overall mean test score, 78 percent of blacks and 65 percent of Hispanics have lower test scores than the common mean. In Table 2, we use the white AFQT distribution as a baseline, and construct AFQT quartile cutoff levels in which one fourth of the white sample falls within each quartile. Table 2 (Panel A, quartile 1) shows that 73 percent of blacks and 59 percent of Hispanics score in the bottom quartile of the white distribution. This pattern continues in the second quartile, with 90 percent of blacks and 82 percent of Hispanics scoring lower than the median score of the white distribution.

Measured ability gaps are substantial and consistent even for the same level of schooling completed at the time of the test. Panels B to G of Table 2 contain the distribution of AFQT using the quartiles of the white sample as the baseline. For each schooling level, minorities are clustered in the lower quartiles of the white distribution. The disparities decline slightly among those with nine or fewer years of completed schools (Panel B). One possible explanation for these disparities is that blacks and Hispanics start school late and are more likely to repeat grades. Gaps in educational attainment between whites and minorities are extensive and influence the measured gap in cognitive ability. Educational disparities are likely to stem from uneven opportunities for advancement in elementary and middle school and from environmental disparities at earlier ages. Phillips, Crouse, and Ralph (1998) demonstrate that black children enter school with lower math and reading skills than their white peers, and that the disparities increase throughout elementary and middle school. Table 3 shows educational attainment at the time of the test for a given age for each demographic group. At age 16, each demographic group's modal attainment is tenth grade (Panels A to F). While 67 percent of white men reach tenth grade or higher, only 51 percent of black men and 50 percent of Hispanic men have reached that level. Similarly, 82 percent of white women reach tenth grade or higher, while only 72 percent of black women and 57 percent of Hispanic women reach the tenth grade. The pattern is similar at upper ages—blacks and Hispanics lag behind their white peers.

Moreover, it has been documented that blacks and Hispanics who are behind in grade level are more likely to drop out of school and are less likely to enter college after completing high school. In our data, the racial-ethnic differences at the time of enrollment are substantial: 21 percent of white men were not enrolled in school in tenth grade, compared to 33 percent of black men and 38 percent of Hispanic men. If selective attrition of the sample is present, low-ability individuals detached from school at the time of the test will likely bias upward estimates of the effect of schooling on minority test scores. Figure 2 presents the AFQT distributions based on final schooling attained. The racial and ethnic disparities found when controlling for schooling level at test date translate to measured ability gaps for those with similar final schooling levels.

2.1 Ethnic Disparities in Environments

Part of the problem in explaining the achievement gap is the lack of a comprehensive dataset that contains a large set of covariates and that would allow us to deal with the complexities of educational attainment. In this regard, the NLSY family background and family environments are much less rich than those in the CNLSY or ECLS. However, the NLSY has information on educational and labor market outcomes, which cannot be observed due to the age of participants in the ECLS and CNLSY. In this section, we present basic evidence of racial disparities in environments within the NLSY data used in our model estimation. Tables 4 and 5 show racial disparities in school quality, environment and family characteristics and school quality measures. Table 4 shows that minorities are more likely to come from a "broken" (or single-parent) home at age 14, have greater number of siblings and have parents with lower levels of education. Half of the black sample came from a single-parent family at age 14, compared to 20 percent of whites and 20 percent of Hispanics. Parental education also differs greatly. While the parents of whites had on average finished high school (12 years of schooling), parents of blacks had on average two years less schooling and parents of Hispanics had on average almost 4 year less of schooling compared with parents of whites. Family income in 1979 presents an important racial and ethnic inequality. The average income of whites was almost twice as large as that of blacks and was 60 percent larger than that of Hispanics.

Table 5 presents school characteristics obtained in the first wave of the NLSY. From this table, we observe racial-ethnic differences in sample means in a few school quality characteristics. There are more books per student in schools attended by whites than schools attended by minorities. Turnover of teachers is more frequent in schools attended by Hispanics. Given the evidence presented earlier, it is not surprising that blacks and Hispanics attend schools with a greater proportion of students classified as disadvantaged and with higher dropout rates. However, there are no significant differences in other measures such as teachers or counselors per student, teachers with graduate degrees or teacher salary. Ferguson (2002b) shows that school factors such as student-teacher ratio are no longer correlated with the school's racial composition. Ferguson documents racial disparities in other school factors that affect educational achievement, such as quality of teaching, class size, teaching expectations and student motivation. Although the NLSY does not provide comprehensive measures of school quality, which can temper the estimated effect of schooling on cognitive ability, by including these variables in the models we can capture systematic racial discrepancies related to unobserved factors affecting educational attainment.[3]

The first methodology presented in this paper uses a unique feature of the NLSY to approximate the effects of schooling on measured ability. This model controls for endogeneity of schooling decisions and identify the effect of schooling in measured ability. Given the important racial-ethnic differences in family and school inputs, and other environments, we explore the effect of this factors on test scores. Differential effects of this factors on test scores across racial-ethnic groups will help identify which policies are more effective in reducing racial gaps in test scores.

3 The Effect of Schooling on Test Scores

In this section, we present a methodology for identifying the causal effect of schooling on test scores.[4] Previous research indicates a sizable effect of schooling on measured ability and identifies the problem of reverse causality, in which test scores influence educational attainment.[5] Various identification techniques have been proposed to estimate the effect of schooling on test scores.

[3]See Ferguson (2002b) for an extensive review of the importance of school quality on racial disparities in educational achievement.

[4]Ceci (1991) surveyed the literature on education's effect on cognitive ability.

[5]Neal and Johnson (1996), Winship and Korenman (1997) and Hansen, Heckman, and Mullen (2004) calculate sizable effects of schooling on test scores. These authors calculate that an additional year of schooling increases AFQT by as many as 2 to 5.4 percentage points.

Hansen, Heckman, and Mullen (2004) show that these strategies are ambiguous as to what constitutes cognitive ability, and that previous methodologies do not identify the causal effect of schooling on test scores.[6] The major shortcoming of the literature is that it fails to account for selection in educational attainment caused by differences in latent ability.

Schooling is likely to have a different effect within different demographic groups. Blacks and Hispanics are more likely than whites to drop out of high school, and those enrolled are more likely to be in a lower schooling level than their white peers of the same age. Socioeconomic status, school quality and family environment differences among whites and minorities have been extensively documented,[7] but a considerable test-score gap still remains after controlling for these factors. Controlling for disparities in schooling and age at the time of testing is crucial in understanding why the scholastic ability of minorities remains behind that of their white counterparts.

To address these issues, we employ a model proposed by Hansen, Heckman, and Mullen (2004) to estimate the racial-ethnic gap in test scores. Their methodology accounts for the endogeneity of schooling by postulating that both schooling and test scores are generated by a common latent ability. The identification strategy is based on differences in final schooling attainment for individuals with the same schooling levels at the time of the test. The model exploits differences in schooling at the test date and in final schooling to identify the effects of schooling on test scores for a given level of ability.

3.1 Identification Strategy

Following Hansen, Heckman, and Mullen (2004), we express measured ability as manifestations of schooling s, latent ability f, and other covariates X:

$$T\left(s,x\right)=\mu\left(s,x\right)+\lambda\left(s,x\right)f+\varepsilon\left(s\right)$$

where $T\left(s,x\right)$ is the test score of individuals with s years of schooling and other factors that influence test scores x. The term $\mu\left(s,x\right)$ is the effect of schooling (and covariates) on test scores, which is

[6] See Winship and Korenman (1997) and Hansen, Heckman, and Mullen (2004) for a review of shortcomings of other identification techniques.

[7] Ferguson (2002a,b) reviews evidence on differences in family inputs and environments.

uniform across latent ability levels. Latent ability f captures the portion of cognitive ability, which is fixed at early ages and cannot be influenced by further interventions and environments. IQ testing is intended to capture f. Therefore, we assume $f \perp\!\!\!\perp X$. Assume that f and $\varepsilon(s)$ have zero means, $\varepsilon(s) \perp\!\!\!\perp X$. The term $\lambda(s,x)$ is the effect of schooling (and other covariates) on transforming latent ability. As argued before, the schooling effect on test scores will likely vary for different levels of latent ability. This is captured by discrepancies in $\lambda(s,x)$.

In the absence of a measure of latent ability f for a given test score $T(s)$, we cannot identify the effect of schooling on test scores. To demonstrate this, consider estimating the effect of schooling:

$$E[T(s) \mid S=s] = \mu(s) + \lambda(s) E[f \mid S=s] + E[\varepsilon(s) \mid S=s]$$

where for simplicity we omit the dependency on X. Since $E[f \mid S=s]$, latent ability determines schooling attainment and influences test scores, hence we cannot identify the causal effect of schooling on test scores. In the case $E[\varepsilon(s) \mid S=s] \neq 0$, we cannot identify the causal effect of schooling on test scores. However, if test scores are not used to make schooling decisions, as in our case, then $E[\varepsilon(s) \mid S=s] = 0$.

The identification strategy in this model accounts for the endogeneity of completed schooling. For individuals with the same schooling at the time of the test, differences in educational attainment later in life allow us to account for differences in latent ability.[8] Therefore, we distinguish between these two schooling measures, final schooling attained, defined as S, and schooling at the time of the test, defined as S_T. Hence, test scores depend on schooling at the time of the test, latent ability and other factors

$$T(S_T, X) = \mu(S_T, X) + \lambda(S_T, X) f + \varepsilon(S_T) \tag{1}$$

where $\mu(S_T, X)$ and $\lambda(S_T, X)$ dependency on X is not necessarily linear. Test scores conditional

[8]Using a feature of the NLSY where individuals tested in 1980 at ages 16 to 23 have different levels of completed schooling. At 1980, some individuals had not completed their high-school education or had dropped out, others had finished high school and others pursued postsecondary studies. In 2000, we are able to observe the final educational attainment of these individuals.

on final schooling, schooling at test date and X can be expressed as

$$E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] = \mu\left(s_T, X\right) + \lambda\left(s_T, X\right) E\left[f \mid S_T = s_T, S = s, X\right]$$
$$+ E\left[\varepsilon\left(S_T\right) \mid S_T = s_T, S = s, X\right]$$

If individuals consider only final schooling when making schooling decisions and sampling is random across ages, S_T will be random with respect to f, conditional on S. Then, $E\left[f \mid S_T = s_T, S = s, X\right] = E\left[f \mid S = s\right]$.[9] Since $E\left[\varepsilon\left(S_T\right) \mid S_T = s_T, S = s, X\right] = 0$ the last equation simplifies to[10]

$$E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] = \mu\left(s_T, X\right) + \lambda\left(s_T, X\right) E\left[f \mid S = s\right]. \tag{2}$$

For two final schooling levels $s, s\prime \geq s_T$, we can estimate the difference in expected test scores for individuals with the same schooling level at the date of the test.

$$E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T, S = s', X\right]$$
$$= \lambda\left(s_T, X\right)\left[E\left[f \mid S = s\right] - E\left[f \mid S = s'\right]\right] \; .$$

Since we do not observe f, we can only identify $\lambda\left(s_T, X\right)$ up to a scale. Given two values of schooling at test date $s_T, s_T' \leq s, s'$ and $\lambda\left(s_T', x\right) \neq 0$, we can form the ratio

$$\frac{E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T, S = s', X\right]}{E\left[T\left(S_T\right) \mid S_T = s_T', S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T', S = s', X\right]} = \frac{\lambda\left(s_T, X\right)}{\lambda\left(s_T', X\right)}.$$

Define the maximum schooling levels as $\bar{S}$, from the previous equation we are able to identify $\lambda\left(s_T, X\right)$ for $s_T = 1, \ldots, \bar{S} - 1$ up to a normalization.[11] $\lambda\left(\bar{S}, X\right)$ is not identified. With these values we can identify $\mu(s_T, X)$ for $s_T = 1, \ldots, \bar{S} - .1$ from the test equation. In addition, the terms $E\left[f \mid S_T = s_T\right]$and $E\left[f \mid S = s\right]$ are identified for all values of s and s_T (see Hansen, Heckman,

[9]See Hansen, Heckman, and Mullen (2004) for an explicit account of assumptions. They do not explicitly derive the model when test scores depend on covariates X. Refer to part B of Appendix B in this paper for a detailed derivation of the model with explicit accounting for X.

[10]Using the fact that test scores do not influence schooling decisions $E\left[\varepsilon\left(S_T\right) \mid S_T = s_T, X\right] = E\left[E\left[\varepsilon\left(S_T\right) \mid S_T = s_T, S = s, X\right] \mid S_T = s_T\right] = 0$, and given that $\varepsilon \perp\!\!\!\perp X$.

[11]We cannot identify $\lambda(\bar{S}_T, X)$ because there is one possible value, $\bar{S}_T = \bar{S}$.

and Mullen, 2004, and Appendix A for the complete derivation of the identification strategy of the model).

In the next section we present estimates of the model, which extends the analysis for Hansen, Heckman, and Mullen (2004) in two important aspects. First, it estimates the effect of schooling on measured ability for white, black and Hispanic men and women. Second, it estimates the effect of family, school inputs and other environments on test scores.

4 Empirical Implementation

This section estimates the model presented in the previous section. First, the model is estimated without including the effect of covariates X on test scores, which constitutes our baseline model. Second, we extend the model to control for family inputs and environments that affect test scores. In the next section, we present comparative results showing the effect of schooling on test scores from a nonparametric structural model.

4.1 Estimates of the Baseline Model

First, we report nonparametric estimates for the control function—which constitutes our baseline model— without including other covariates that could affect test scores.[12] Following the previous analysis

$$T\left(s_T\right) = \mu\left(s_T\right) + \lambda\left(s_T\right) f + \varepsilon\left(s_T\right).$$

where $T\left(s_T\right)$ are AFQT scores at schooling level s_T. Figures 3 plot estimates of $\mu\left(s_T\right)$ and test score sample means for each demographic group and five schooling levels.[13] $\mu\left(s_T\right)$ corresponds to the effect of schooling on test scores, independent of latent ability The results confirm previous research about the magnitude of the racial gap in test scores. For the same schooling levels, whites significantly outperform minorities at each schooling level and Hispanics outperform blacks. We can see that the test scores gap widens with schooling. From the figure, we observe sharp contrasts in the effect of schooling across race and gender. Hispanic men and white women seem to benefit more

[12]The first part of Appendix B describes in detail the econometric procedure.

[13]As explained in section 3, $\mu\left(\bar{s}_T\right)$ is not identifiable since we cannot identify $\lambda\left(\bar{s}_T\right)$.

from postsecondary education than any other group. The term $\mu(s_T)$ is plotted against the raw data sample means $E[T(s_T) \mid S_T = s_T]$. The term $\mu(s_T)$ is plotted against the raw data sample means $E[T(s_T) \mid S_T = s_T]$. The difference between the two lines is given by $\lambda(s_T) E[f \mid S_T = s_T]$,[14] which represents the proportion of test scores explained by latent ability. Latent ability explains a large part of test scores for black and white men who attend some college. The percentage of sample means $E[T(s_T) \mid S_T = s_T]$ not explained by $\mu(s_T)$ reached 19 percent for black men and 10 percent for white men, compared with 3 percent for Hispanic men (Panel A). Analogously, latent ability represents 21 percent of test scores for black women and 16 percent for Hispanic women who attended postsecondary schooling, in contrast with 6 percent of white women (Panel B).

The model identifies the causal effect of schooling, for two levels of schooling at test date $s_T < s'_T$, given by $\mu(s'_T) - \mu(s_T)$. Table 6, Panel B shows the causal effect of an additional year of schooling. In general, schooling has a greater effect on whites than blacks, but not Hispanic men. The first row of Panel B shows the causal effect of moving from 9th grade or lower to 10th grade—this produces the greatest effect on AFQT scores (up to 13.29 AFQT points) for each group. This is not surprising since it is the effect of more than one year of schooling. The effect of an additional year of schooling for those in 10th grade is substantially higher for men than for women of any ethnicity. White men increase 4.94 AFQT points with an 11th grade education compared to 1.92 AFQT points for white women (Panel B second row). Black men score 4.1 AFQT points more than black women for an additional year of schooling, and Hispanic men score 2.8 AFQT points more than Hispanic women from 10th to 11th grade. High-school completion reverses gender benefits from schooling for black women (Panel B, third row). High school graduation and especially postsecondary education have greater benefits for Hispanic men. However, a considerable AFQT score gap persists between Hispanics and white men. Blacks start behind and remain behind. Schooling effects on AFQT scores at each grade are lower compared to schooling effects for whites and Hispanics. These results support the crucial role of accounting for schooling levels at test date to explain racial disparities in scholastic ability and its effects on labor market outcomes.

[14] From the equation

$$E[T(s_T) \mid S_T = s_T] = \mu(s_T) + \lambda(s_T) E[f \mid S_T = s_T]$$

The estimates of factor loadings $\lambda(s_T)$ are presented in Table 7. The estimates represent the effect of schooling on transforming or revealing latent ability. This table shows the unrestricted estimates of factor loadings for different comparison groups. Since we can only identify the factor loadings up to a scale, the estimates are normalized with respect to $\lambda(S_T = 2 = \text{10th grade})$. For $\lambda(S_T = \text{9th grade or less})$ and $\lambda(S_T = \text{11th grade})$, we have six unrestricted estimates and three unrestricted estimates for $\lambda(S_T = \text{High school})$. One factor loading is estimated for $\lambda(S_T = \text{ 13,14 years})$. Comparing the $\lambda(s_T)$ estimates, test scores of whites and Hispanics increase at a lower rate than those for blacks for each additional year of schooling. Schooling plays a larger role in revealing latent ability among blacks, especially for 12th grade and postsecondary education. The effect of schooling, for a given level of latent ability, is given by $\mu(s'_T) - \mu(s_T) + [\lambda(s'_T) - \lambda(s_T)] f$. The largest effect of schooling attributable to $\lambda(s'_T) - \lambda(s_T)$ for black men originated from high-school graduation, not from postsecondary education. Black and Hispanic women experience a greater schooling effect of attending postsecondary education.

The control functions estimates, which are expected values of latent ability conditional at each schooling level are presented in Table 8. Panel A of Table 8 shows $E[f \mid S_T]$ for the six schooling levels at test date. As expected, latent ability is positively correlated with schooling. However, the term $E[f \mid S_T]$ does not have a clear interpretation since the AFQT test was administered with a sample rule not directly related with school completed at test date. Panel B of Table 8 presents $E[f \mid S]$ for the four levels of final schooling. As anticipated, expected latent ability increases with schooling for each demographic group. High-school dropouts are at the bottom of the ability distribution compared to college graduates for each group. The underlying latent ability for each schooling level is consistent across demographic groups. The mean ability is reflected as postsecondary education within each group. For $s_T \leq s < s'$, we can interpret $\lambda(s_T)[E[f \mid S = s', S_T = s_T] - E[f \mid S = s, S_T = s_T]]$ as the difference in test scores for two individuals who were tested at the same schooling level but achieved different final levels of education. The fact that $\lambda(s_T)$ is decreasing with schooling for whites indicates that test scores magnify differences in ability at lower schooling levels and diminish differences in ability at higher schooling levels. For blacks, the theory is reversed: since $\lambda(s_T)$ increases with schooling, test scores will show greater differences due to ability differentials at higher schooling levels. These results indicate that test scores reflect ability in different ways for different ethnic groups. When comparing two men who are high-school graduates at the time of the test and one continues to college, the test scores will reflect damped ability differentials for whites and increased differentials for blacks. This result is particularly important when comparing test scores between ethnic groups and controlling for ability differentials in wage equations. However, it should be interpreted with caution given that the underlying ability distributions differ for each demographic group considered.

In summary, the model allows us to identify two ways in which schooling affects test scores. First, test scores increase with schooling, independent of latent ability. The effect of schooling on measured ability is greater for whites than blacks but not for Hispanic men. Second, we can measure the effect of schooling on revealing or transforming ability as shown by test scores. In this section, we estimate a substantial causal effect of schooling on test scores by controlling for the reverse causation of education and AFQT. This contradicts the evidence presented by Herrnstein and Murray (1994) about the immutability of their ability measure (g or AFQT). Also, we have recreated the racial-ethnic gap found in test scores and gone a step further in our understanding of the scholastic achievement gap since we can identify the components of the test-score gap. In Table 9 we observe that a large portion of the ethnic test-score gap is due to disparities in $\mu(s_T)$. Since this term defines a mean, it likely captures the effect of other determinants of test scores. Given this distinction, controlling for other covariates will likely affect $\mu(s_T)$. Changes in the composition of $\mu(s_T)$ will increase our understanding of the causes of the racial gap in test scores. Next we present extensions of the baseline model to incorporate covariates, such as school quality and family environments, that have been proposed as determinants of achievement test scores by previous literature.

4.2 School and Family Determinants of Test Scores

Understanding the determinants of test scores is essential to identify the sources of racial-ethnic differences in cognitive ability. Recent research suggests that the dynamics of skill formation along the life cycle is cumulative: early investments in skill and favorable environments are essential in fostering cognitive ability later on. While previous literature, surveyed by Ceci (1991), suggests that cognitive skill is influenced by family and school inputs, there has not been a consensus of what are the main determinants of cognitive ability; there are often contrasting effects of family environments and school variables on test scores. The main problem the literature has faced is that measures of cognitive ability and environments are not generally available through the life cycle, which hinder deriving a causal relationship between them. To overcome these limitations, previous studies propose different models that, while they do not adequately address the reverse causality of environments, schooling and cognitive ability, do provide correlational evidence of test scores and environments.

In contrast with previous literature, our identification strategy allow us to address the problem of the reverse causality of schooling and ability which allow us to control for the effect of covariates on test scores. AFQT scores, measured later in life ages 16 to 23, are likely to capture the effect of environmental variables and school inputs. This methodology explores the effect of covariates on test scores which allows us to identify the efficacy of skill investment policies in different racial groups later in life. As in previous literature, our main obstacle in estimating the determinants of test scores is data limitation. The methodology requires that we observe test scores for individuals with the same schooling levels at test date and achieve different levels of final schooling for each demographic group. Hence, due to small sample sizes, we pooled males and females of each ethnicity to estimate the effect of test score determinants. In addition, the number of covariates that we can include simultaneously is limited. In order to implement this strategy, in the rest of the section we allow covariates to affect test scores as defined in equation (1)

$$T\left(S_T, X\right) = \mu\left(S_T, X\right) + \lambda\left(S_T, X\right) f + \varepsilon\left(S_T\right).$$

where covariates X affect the component of the test score independent of latent ability, $\mu\left(S_T, X\right)$ and the way latent ability is revealed or transformed, captured by the term $\lambda\left(S_T, X\right)$. In all cases we include a gender dummy variable to account for systematic gender determinants of test scores within each race. Next we present the estimation results from the model including school quality and family background variables.[15]

[15]Our empirical methodology is explained in Part 2 of Appendix B.

4.2.1 Family Environments

We include in vector X environments and family background variables summarized in Table 4. In Figure 4 we report estimates of the component $\mu(S_T, X)$ for whites, blacks and Hispanics, including family background variables. First, estimates of the baseline model for each level of schooling are depicted in the far left. This corresponds to estimates excluding environment covariates. Family background variables are included in X one at the time in the next five columns in addition to a gender dummy. We include family income, broken home, number of siblings and mother's and father's education. In the far right column, the five family environments are included.[16] From the three panels we observed that the inclusion of family covariates do not significantly affect the estimates. AFQT as measured at 16 to 23 years of age is not affected by family environments. Schooling at the time of the test is the main determinant of test scores. For every race, the estimates for 9th grade or less, 10th grade, 11th grade and high school do not differ significantly from baseline estimates. The same is true for whites and Hispanics with 13-14 years of schooling. For blacks with postsecondary education, the estimates controlling for family environments depart from baseline estimates. In particular, mother's and father's education levels increase the effect of postsecondary schooling on test scores for blacks.

4.2.2 School Quality

Although there is no consensus about whether school inputs are important determinants of cognitive ability, the substantial effects of schooling on test scores suggest a need for closer analysis of school inputs. Previous studies do not find a significant effect of school quality on achievement test scores or earnings.[17] The survey of schooling characteristics in the NLSY, conducted in 1980 and restricted to individuals who are 17 years old, presents high non-response rates. Previous Table 5 summarizes school quality characteristics considered in our estimates. For whites, school variables in our sample are available for 69 to 78 percent of the original sample. For blacks, school characteristics are present from 53 to 65 percent of the sample. In turn, 59 to 66 percent of Hispanics of our sample present

[16] For simplicity we do not show results from living in the south at age 14 and urban at age 14. These estimates do not differ significantly from the baseline model for each group.

[17] See Betts (1995); Boozer, Krueger, and Wolkon (1992); Hanushek (2004); Heckman, Layne-Farrar, and Todd (1996); Heckman and Neal (1996).

school characteristics.[18] Racial-ethnic differences in enrollment and high-school dropout rates could lead to selection attrition in measures of school characteristics. High-school dropouts are more likely to lack information on school characteristics and have low levels of school inputs. There is not evidence that NLSY data present systematic missing school information from high school dropouts. The magnitude and direction of a potential bias caused for selective attrition on test scores is not clear. Attrition from the sample of the less able individuals will overestimate the impact of schooling on test scores if poor school inputs significantly impact dropping out of school.

Figure 5 presents estimates of $\mu(S_T, X)$ including measures of school quality for each schooling level at the time of the test. Due to small sample size, we cannot include several covariates at the same time. The first column represents baseline estimates (excluding covariates), restricted for the sample that contains information for all four school characteristics: teacher turnover, books per student, daily percentage attendance in school and faculty per student (pupil-teacher ratio).[19] The next four columns include in the estimation each covariate, one at the time. The last column includes estimates of $\mu(S_T, X)$,including the four school characteristics. From Figure 5 we observe that school quality does not change the baseline estimates for the white sample for every schooling level. For blacks and Hispanics we obtain a similar result for schooling levels below 13-14 years. Similar results are obtained if we include other school characteristics such as teacher education and teacher salary. For 13-14 years of schooling the estimates of minorities, in particular Hispanics, deviate from the baseline estimates. However, this can be attributed to small sample sizes (Table A1).

Since AFQT and the survey of schooling are contemporaneous, it is surprising that traditional measures of school quality do not have a sizable effect on achievement test scores. Schooling attainment is the main determinant of test scores. There are two important considerations when interpreting previous as well as our own results. First, traditional measures of school quality—such as teacher-pupil ratios, class size, teacher salary—cannot accurately reflect racial differences in school inputs. A reduction of racial-ethnic gaps in school characteristics has been documented

[18] Books per student present the smallest sample sizes, which contains the smallest number of observations for every ethnicity.

[19] A gender dummy variable is included in every case. Each school characteristic is available for almost the same subsample of individuals. We do not loose many observations limiting the sample to non-missing values for all four characteristics.

in the literature. Ferguson (2002b) present a good survey of school characteristics and black-white school achievement gap. Boozer, Krueger, and Wolkon (1992) find that pupil-teacher ratios are similar for blacks and whites and higher for Hispanics. This reduction was fueled from federal and local government policies to increase school quality and reduce racial segregation (Hanushek, 1989, 2004; Ferguson, 2002b). Since these policies were directly targeted to improve such indicators, racial differences in school quality cannot be reflected in traditional school indicators such as pupil-teacher ratios.

Although the identification strategy is appealing, the model cannot be easily extrapolated to other datasets with different schooling structures. In addition, since the model relies on variations of final schooling levels and schooling at test date, the model has limitations in the number of controls that can be included for small sample sizes. To overcome this issue we estimate a structural model that controlling for unobserved heterogeneity allows the identification of the distribution of latent ability f. With this model we can condition covariates that affect test scores, such as family environments, school factors and local labor market variables, which may influence schooling choices.

5 Structural Model Estimation

This section presents the main features of a semi-parametric structural model, extended from standard factor analysis.[20] We present empirical estimates of the model and compared them with our estimates of the previous section. In addition, we further explore the effect of school quality on test scores. This methodology does not rely on a special randomization structure of the NLSY ASVAB data and it can be applied to other datasets. This model requires that two or more tests are observed for the same individual. Define $T_K(S_T)$ as the score on the k^{th} test ($k = 1, ..., 4$), defined as the four ASVAB components of the AFQT score[21], for individuals with schooling level S_T at the time of the test. Test scores are influenced by latent ability and other observable determinants

[20] The complete specification of the model can be found in Hansen, Heckman, and Mullen (2004). Our empirical specification is based on their Appendix C, adapted for six different demographic groups and four final schooling choices.

[21] We considered the following ASVAB tests: word knowledge, paragraph comprehension, arithmetic reasoning and math knowledge. The tests are standardized within each demographic group.

defined by $x\,(s_T)$:

$$T_k(s_T) = \mu(s_T, x) + \lambda_k(s_T) f + \varepsilon_k(s_T). \tag{3}$$

where we assume a linear term $\mu(s_T, x) = X(s_T)\beta_k(s_T)$.We include the following environments as determinants of test scores, $X(s_T)$: urban status, broken home and southern residence at age 14, number of siblings and family income in 1979, mother's and father's education and age in December 1980. In addition a dummy variable was included for those individuals who achieved 9th grade or lower at the time of the test. We assume $\varepsilon_k(s_T) \sim N(0, \sigma_k(s_T)^2)$. We also take into account the ceiling effects for individuals who achieve the maximum score on the test. In this case, the observed test score is given by

$$T_k = \begin{cases} T_k^*(s_T) & \text{if } T_k^*(s_T) < c_k \\ c_k & \text{if } T_k^*(s_T) > c_k \end{cases} \quad \text{for each } S_T = 1, ...\bar{s}_T \text{ and } k = 1, ..., 4.$$

where c_k is the maximum score on the k^{th} test.

To account for endogeneity of schooling, we estimate a semiparametric model of schooling choice. An individual selects schooling that maximizes her utility

$$S = \arg\max{}_s \left\{V(s)\right\}_{s=1}^{s=\bar{s}}$$

where $V(s)$ is the utility associated with final schooling level s.[22] We assume a functional form for utility that is linear in determinants $Z(s)$ and latent ability f:[23]

$$V(s) = Z(s)\gamma(s) + \alpha(s) f + u(s), \qquad s = 1, ..., \bar{s}$$

where the $Z(s)$ vector includes variables that affect schooling decisions, such as family environments, socioeconomic status at early ages and costs of attending school. The choice equations include the same family environments as the test equation.[24] They also include measures of opportunity costs

[22]As before, $S = 1$ represents high school dropout and GEDs, $S = 2$ represents high school graduates, $S = 3$ represents some college and associate degrees, and finally $S = 4$ represents college graduates.

[23]This assumption is taken for simplicity; utility can have a nonlinear specification.

[24]The covariates include urban status, broken home and south residence at age 14, number of siblings and family income in 1979, mother's and father's education and age.

that affect schooling decisions, such as local tuition for two and four year colleges, local wages and local unemployment rates constructed for each schooling level.[25] The error term is assumed to have a standard normal distribution, and $Cov(u(s), f) = 0$. This model of choice is then adjoined to the system of test scores. For simplicity we omit the steps of identification of the model, which can be found in Hansen, Heckman, and Mullen (2004). Next, we present estimation results for each of the six demographic groups. One important feature of the structural model is that it allows us to estimate the effect of ability on test scores at various levels of the ability distribution. These estimates supplement the control function model of the previous section since they can be applied to datasets that do not have the schooling structure of the NLSY. In our NLSY sample context, the empirical implementation of the structural model allow us to include more covariates than our control function approach.

5.1 Comparative Estimates of Control Function and Structural Model

In this section we first present comparative results of the structural estimates and the baseline control function model. Second, we estimate marginal effects on AFQT scores of latent ability interventions. Finally, we present estimates of the test score for different percentiles of the ability distribution.

We begin with the analysis estimates of $\mu(S_T, X)$ for both models are presented in Figure 6. Structural model estimates are plotted against the baseline estimates of the control function $\mu(S_T)$ and sample means for each schooling level. For each demographic group, the estimates of both models are in close agreement for schooling levels less than 12th grade but there are differences for postsecondary levels. Compared with the baseline model, the structural estimates show a stronger effect of attending postsecondary education on test scores for black and white men. For women, we have similar results: the structural model shows larger schooling effects on AFQT scores. Both estimates are in close agreement with the sample means for each schooling level below high school (12 years). The unexplained portion of the mean test scores—which includes the latent ability component—is smaller for estimates of the structural model for 13-14 years of schooling: for white men, the unexplained portion of test scores is 4 percent of the structural model estimates compared

[25]A extensive discussion of covariates is included in Appendix A.

with 12 percent of the baseline model estimates. For blacks, it represents 12 percent of test scores in the structural specification and 22 percent in the baseline model—contrary to an increase from 4 to 9 percent for Hispanic men. For white women, the portion of test scores explained by latent ability is smaller (3 percent). Unlike with the control function approach, we can estimate $\mu(S_T, X)$ for $\bar{S}_T$. The estimates show a significant departure from sample means for 15 years or more of schooling, especially for blacks in which the unexplained portion of test scores is 14 percent. The latent ability component of test scores is larger for postsecondary schooling levels. Unfortunately, in this case we cannot estimate for Hispanics due to small sample sizes.

The effect of schooling on test scores, defined by $\mu(S_T', X) - \mu(S_T, X)$, is presented in Table 10. This table presents structural estimates similar to those of the control function model for schooling levels high school level and below: an additional year of schooling has a greater effect for whites than for minorities. As expected, the estimated effects of attending some college are larger than the baseline model for blacks and whites. For black men, there is an increase of 6.81 AFQT points with college attendance and 10.12 AFQT points for white men (Panel A row 4). The largest effect of attending college is for black men with 11.91 AFQT points (Panel A, row 5). Black and white women have a larger effect for attending 13-14 years of school, but this effect does not persist to attendance of 15 years or more (Panel B, row 5). The estimates of factor loadings $\lambda(S_T)$ from control function and structural models are in close agreement. Figure 7 shows the ratios of factor loadings normalized at 10 years of schooling for each demographic group. The figure shows a similar pattern for the factor loadings for males: they decrease with schooling for whites and increase for blacks (Panels A and B respectively). For Hispanics, structural estimates of factor loadings increase with schooling and diverge from the control function estimates for 13-14 years of schooling (Panel C). A similar result is found for females (Figure 8): factor loadings are decreasing with schooling for whites and increasing for blacks.

We identify the distribution of latent ability f and explore its relationship with achievement test scores. With the distribution of latent ability in hand, we can estimate the schooling effects on test scores at different latent ability levels. Figures 9 and 10 present expected test scores at the 2.5th, 50th and 97.5th percentiles of the latent ability distribution by schooling at the time of the test. As expected, test scores are higher when conditioning at higher ability percentiles for every schooling level (Panels A to C). For a given demographic group, the effect of schooling differs substantially as we condition at different ability levels.[26] We observe sharp differences between whites and minorities. For whites, the effect of schooling on test scores declines sharply (except for the effect of 9th to 10th grade). Comparing Panels A to C, for white men at the bottom 2.5th percentile of the distribution, attending 13-14 years of schooling increase AFQT scores by 14.34 AFQT points (compared with 12 years of school), in contrast with 9.96 AFQT points for white men at the median of the distribution, and 6.64 AFQT points for white men at the 97.5th percentile. In contrast with whites, the causal effect of schooling for minorities *increases* when we condition for higher ability levels. This is a striking result, for even black men at the top of the ability distribution show a significant increase of AFQT scores— 9.8 AFQT points for attending 15 or more years of schooling, compared with 0.03 AFQT points for white men. We have to be cautious with this result since our sample has only 37 black men who attended 15 years or more of schooling at the time of the test. It is evident that the effect of schooling on test scores differs greatly between ethnic groups at different latent ability levels. As in the baseline model estimates, whites have a greater effect of schooling on test scores, but *only* for lower ability levels.

Table 11 shows the effect on test scores of an intervention that increases latent ability by one standard deviation for different demographic groups (under Model 1). The results confirm the insights from the baseline model. The marginal effects on AFQT are similar between races at 12 years of schooling (high-school graduation). A one-standard-deviation increase in latent ability has similar effects on AFQT for different ethnic groups below high-school level. For higher schooling levels, minorities benefit more from an increase in ability even at higher schooling levels.. The effect on test scores of the increase in latent ability diminishes with schooling, but on a larger scale for white men and women.[27]

Estimates of both models are in close agreement. The structural estimates are in closer agreement with data for schooling levels above high school. This methodology allows us to assess the impact of school quality at various levels of the latent ability distribution.

[26] Recall that the causal effect is difference of test scores when controlling for a fixed latent ability level.

[27] Previous research has confirmed this result with data form ECLS and CNLSY. See Ferguson (2002b), Phillips, Brooks-Gunn, Duncan, Klebanov, and Crane (1998) and Carneiro, Heckman, and Masterov (2005).

5.2 Effect of School Quality on Test Scores

In this section we extend the structural model to control for school inputs that affect test scores. Our estimates show large effects of schooling on cognitive ability, even controlling for family background characteristics. Our results from the control function model do not play an important role in school characteristics. We propose two methods to analyze the role of school inputs on achievement test scores. We compare this two methods with estimates of our structural baseline model that does not include school quality (hereafter Model 1) represented by equations (3). First, school quality enter linearly in the test equation

$$T_k(s_T) = X(s_T)\beta_k(s_T) + Q(s_T)\delta_k(s_T) + \lambda_k(s_T)f + \varepsilon_k(s_T), \tag{4}$$

where $Q(s_T)$ are school quality characteristics for individuals at school level s_T at the time of the test (hereafter we refer to this specification as Model 2). We assume a linear specification for schooling characteristics in the utility function

$$V(s) = z(s)\gamma(s) + Q\theta(s) + \alpha(s)f + u(s), \qquad s = 1, ..., 4.$$

The second specification estimate Model 1 for subsamples of low and high school quality school quality (referred as Model 3). For each demographic group, we identify individuals above and below the mean of each school characteristic for each demographic group. This distinction defines low and a high quality subgroups. Systematic difference in test scores for these subgroups are indicative of a possible influence of school characteristics on test scores. However, since individuals do not randomly assigned to each subgroup, we cannot derive a causal implication from this methodology. Factors affect test scores—such as omitted school characteristics—could affect school quality and then our subsample definition. In this case, we define test scores for each subsample as

$$T_k(S_T, D_Q) = X(S_T, D_Q)\beta_k(S_T, D_Q) + \lambda_k(S_T, D_Q)f + \varepsilon_k(S_T), \tag{5}$$

where D_Q is a dummy variable if individual attended a low quality school (hereafter Model 3).

Estimates are obtain for each subsample by same methodology as Model 1. The model is estimated for subgroups above and below each school variable.

Figures 11-13 presents estimates of test scores for white, black and Hispanic males respectively. Test scores are estimated for fixed levels of the latent ability distribution (25th, 50th and 75th percentiles). Baseline estimates (Model 1) are presented in the first column for each schooling level. In the next five columns, test scores estimates including each school variable linearly in the test equation: annual salary for a new teacher (Q1), teacher turnover (Q2), percentage of faculty with MA or PhD (Q3), book per student (Q4) and faculty per student (Q5).[28]

The three panels of Figure 11 show a similar story for whites. From each column of Model 2, linearly including school variables does not affect our baseline estimates in a significant way: differences are below 1 AFQT point in most cases and below 2 AFQT points in all cases (except for Q1 and Q4 for 15 years or more of schooling at 25th percentile). For white males and females differences in test scores are not significative and negligible[29]. The only pattern from the estimates is a positive effect of 1 AFQT on average for white males and 1.5 AFQT point for females with 15 years of schooling or more. In any case, school quality explain little of the growth in test scores for whites. Test score estimates of low and high quality subsamples (Model 3) present larger differences with baselines estimates than Model 2 in some cases. Differentials of the low quality sample and baseline model are small for white males and females. The estimates of the high school quality sample are larger than the baselines estimates in most cases. However, this estimates are not higher than 2 AFQT points. The only case where there is a clear pattern is for 15 years or more of schooling, where differences in test scores are as large as 4.47 AFQT points for Q1 to Q4 (for the 50th percentile). Surprisingly, for Q5, test scores are consistently *lower* than baseline estimates for 15 years of schooling and higher for low schooling levels (a 25th, 50th and 75th percentiles).

For minorities in figures 12 and 13, comparison of Model 2 and the baseline estimates yield similar results. School quality explain little of growth of test scores among minorities. Due to small sample sizes we cannot estimate test scores including school quality characteristics for black and Hispanics for 15 years or more of schooling (males and females). For black males, 50th and 75th

[28]Estimates of daily attendance to school are not available for structural model.

[29]Point estimates for each schooling level for males and females are available from the authors.

percentile test scores differentials are small. At 25th percentile of the latent ability distribution, test scores are lower when controlling for school quality for black men with 13-14 years of schooling. For Hispanics, estimates of Model 2 are larger for those in 9th grade and lower for those with 13-14 years of more than baseline estimates. We have to be cautions with this results, standard errors for minorities, especially Hispanics, are large. Standard errors for schooling levels above 12th grade due to small sample sizes. For minorities males, school quality estimates are not significantly different than for the overall sample.[30]

Table 12 present point estimates of test scores at 2.5th, 50th and 97th percentile including faculty per student for the three models (4). Comparison of Model 2 and baseline. For white males, the largest test score differential for males and females, at the median of the latent ability distribution (50th percentile), is 1.09 AFQT points for those with 11 years of schooling when we linearly control for faculty per student. Larger differential are found at the end of the distribution. For blacks we obtain a similar result: the test score difference is -1.49 AFQT points. For Hispanics score differentials are larger when controlling for faculty per student for those in 9th grade or lower and smaller for those with 13-14 years of schooling. This reduces the effect of schooling on test scores (-3.77 and 3.19). At the 97.5th and the 2.5th percentiles test scores differentials are larger than at the median of the distribution. For white males at the top 97.5th percentile, the score differential are small and lower than 2 AFQT points, at the bottom 2.5th percentile the largest score differential is 3.19 AFQT points. Table 13 show the same results for females. School quality does not play an important role in explaining test scores. In summary, this results show that school inputs explain little of the differential growth in test scores among blacks, whites and Hispanics.

6 Conclusions

This paper deepens our understanding of the effects of schooling on achievement test scores and questions the importance that has been attributed to inherent ability differences. By controlling for the endogeneity of schooling we identify the causal effect of schooling on test scores. Schooling has sizable effects on measured ability. The estimates show a large effect of schooling on test scores,

[30] For Hispanic Males and Females there were not enough observations for teacher turnover (Q2).

especially among whites. This indicates that policies targeted at improving schooling attainment might prove ineffective and might widen racial-ethnic disparities in achievement test scores. Note that schooling effects on manifest ability differ at different levels of latent ability. Minorities with higher levels of latent ability have a larger gain in test scores for the same increase in schooling. The results indicate that an exogenous increase in latent ability benefits minorities more than whites for higher ability levels. Our estimates from both models are in close agreement. Our results show that racial-ethnic differentials in family and school inputs and other environments does not play an important role ion explaining test scores among whites, blacks and Hispanics. The AFQT scores reflect environmental influences prior to the time of the test, where individuals ranged from 16 to 23 years old. Environmental factors have been influential for a crucial part of the development of cognitive ability when the AFQT measure was taken. The results indicate that interventions aimed at reducing the test score gap by improving conditions early in life are likely be more effective.

A Data Description

This appendix contains details of sample construction as well as brief descriptions of the schooling, family background and family income variables, the Armed Forces Qualification Test (AFQT), local labor market variables, measures of local tuition, and Pell Grant Eligibility.

A.1 Background on the NLSY Data

The National Longitudinal Survey of Youth (NLSY79) comprises three samples that are designed to represent the entire population of youth aged 14 to 21 as of December 31, 1978, and residing in the United States on January 1, 1979. We use the nationally representative cross-section and the set of supplemental samples designed to oversample civilian Hispanics and blacks. The military sample and the sample of economically disadvantaged nonblack/non-Hispanic youths are excluded. Data was collected annually until 1994, then biannually until 2000. NLSY79 collects extensive information on respondents' family background labor market behavior and educational experiences. The survey also includes data on the youth's family and community backgrounds.

Sample Characteristics

Our sample excludes observations of people who were not interviewed some years. This resulted in a loss of about 8 percent of the sample. Another 4 percent of the sample was deleted when individuals who were less than 17 years of age were taken out of the sample. Individuals currently serving in the military, missing values in employment status, or those who are both enrolled in school and working were also removed from the sample. The sample is based on the 1980 cross-section with supplemental information from background and family characteristics from the 1979 interview. After exclusions our sample size is 8,123 (51 percent male). The sample includes individuals from 16 to 23 years old. The ethnic composition is predominately white (4,200 white, 2,429 black, and 1,494 Hispanic).

The racial classifications of between blacks, whites and Hispanics are obtained directly from the NLSY79 guidelines. Hispanics were those who self-identified as Hispanic from the following ethnicities: Mexican-American, Chicano, Mexican, Mexicano; Cuban, Cubano; Puerto Rican, Puertorriqueno, Boriccua; Latino, Other Latin American, Hispano, or Spanish descent. Also classified as Hispanics are persons who did not self-identify as Hispanic but who met the following conditions: those who identified themselves in the ethnic origin categories that included Filipino or Portuguese; those whose householder or householder's spouse reported speaking Spanish at home as a child; and those whose family surname is listed on the Census list of Spanish surnames. Blacks included those for whom race was coded black and ethnic origin was "non-Hispanic" or those whose ethnic origin was coded black, Negro, or Afro-American regardless of race.

A.2 Schooling Choices

A comprehensive highest grade completed (HGC) variable was constructed using several NLSY schooling and enrollment variables. Yearly schooling information is not available for individuals across all years because respondents not enrolled in school since the date of the last interview are not asked their HGC for the year. In light of this, several corrections were employed to maintain consistency across years. The highest grade completed variable at each date is consistent with schooling information gathered in one interview and with the final schooling status. For individuals with missing values for highest grade completed in a given year, the missing values were replaced with the most recent value by looking through the data retrospectively. In cases where this method was insufficient, information on the date of high school certification and college attendance is used to trace the educational path of the individual.

With an accurate HGC variable, we can then sort respondents into educational categories. A high-school dropout is an individual who is not enrolled in school and whose highest grade completed is less than 12 years of school. GED recipients are those who have obtained a high-school equivalence diploma through a General Educational Development certification and do not have postsecondary education. GED recipients are reported in the NLSY to have 12 years of completed schooling. GED holders account for all individuals that we classify as high-school dropouts and have 12 years of completed schooling at the test date. "High school graduate" is defined as those with a high-school diploma through high school certification, with no postsecondary education.

A.3 Socioeconomic Status and Family Structure of the Sample

We create mutually exclusive race-ethnicity categories: white, black, Hispanic, and we obtain family background characteristics for each race-gender group. Family income and background variables include mothers and father's education in 1979, parental family income in 1979 dollars, whether the respondent came from a broken home at age 14 (that is, did not live with both biological parents), number of siblings in the household, and geographic information such as region of residence and urban residence at age 14. In addition we obtained marital status of the individual and region of residence in the US. We impute missing data for parent's education and family income (about 25 percent of the total observations). Predicted values for missing observations were approximated by OLS regression of the nonmissing values on southern residence at age 14, dummy for urban residence at age 14, dummy for broken home status at age 14, number of siblings, and year of birth dummies by race and gender.

A.4 Armed Forces Qualifying Test (AFQT)

The NLSY79 contains the Armed Services Vocational Aptitude Battery (ASVAB), which consists of 10 tests that were developed by the military to predict performance in the armed forces training programs. The battery involves unspeeded achievement tests designed to measure knowledge of general science, arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, coding speed, auto and shop information, mathematics knowledge, mechanical comprehension, and electronics information. The NLSY79 provides raw scores, scaled scores, standard errors, sampling weight, and whether the test was completed under normal or altered testing conditions. The Armed Forces Qualifying Test (AFQT) is the most commonly used combination of tests used by the military for enlistment screening and job assignment by the military. We construct the AFQT from a linear combination of word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests.

A.5 Local Variables

Direct and opportunity costs of attending school affect schooling decisions and must be included in the schooling choice equations. Local wage, unemployment and tuition variables were constructed at the state and Metropolitan Statistical Area (MSA) levels and merged to each individual in NLSY79 sample using the NLSY's Geocode information that provides respondents' geographical location at each interview date. Local wage and local variables were constructed by four education levels in each MSA and attached to each individual at age 17 in the sample according to their education level at a given year. The schooling categories include high-school dropouts, high-school graduates, those with some college or associate degrees, and college graduates. Although the NLSY79 includes local unemployment for each individual in the sample from Bureau of Labor Statistics (BLS) estimates, these estimates do not vary between individuals within an MSA. The NLSY unemployment level does not accurately reflect the economic conditions that individuals with different schooling levels face. Employment opportunities differ greatly for individuals with different skills and schooling decisions and are affected by economic conditions faced by individuals. In our framework, it is particularly important to control for individuals' socioeconomic conditions that affect schooling attainment.

Local Unemployment

Local unemployment rates for each MSA and the four schooling groups were generated from the monthly Current Population Survey (CPS) from 1978 to 2000. Without conditioning on education status, the constructed unemployment rates are highly consistent with the BLS estimates from NLSY79. The sample consists of civilian, non-institutional persons aged 16 to 65 in the labor force. Unemployment rates were calculated for the portion of each state that is not considered part of an MSA (out-of-MSA), and these were assigned to individuals living outside of an MSA.

Local Wage

The local wage was created from the Regional Economic Accounts of the Bureau of Economic Activity (BEA), which provides average wage per job for the MSA and the out-of-MSA portions of the state from 1969 to 2001. The average wage consists of wage and salary disbursements divided by the number of wage and salary jobs (total wage and salary employment). The BEA data does not provide local wage information for different education groups. In order to approximate the wage gap between education groups, we adjust the local wage for the wage gap estimated from the Census data. From the 1980 Census, we calculated the hourly wage from the annual wage and salary income divided by an estimate of the hours worked that year. The wage is estimated by each education group and MSA. In order to calculate the wage that a given individual faced at age 17, when most of the schooling decisions are made, we need annual wage data prior to 1980. Hence, the 1980 Census wage estimate (by education group and MSA) is adjusted with the annual variations in average wage for each MSA from BEA data, to obtain a MSA wage for each year and education group.

Tuition Data and Pell Grant Eligibility

Local tuition at age 17, for two-year and four-year public colleges and including universities, was constructed from annual records on tuition and enrollment from the Higher Education General Information Survey (HEGIS) and the Integrated Postsecondary Education Data System (IPEDS). By matching location with a person's county of residence, we were able to determine the presence of both two- and four-year colleges in an individual's county of residence. Public colleges were divided into two- and four-year programs, and a weighted average of tuition was generated for each county (college enrollment was used for weighting). This process was repeated at the state level.

The tuition data was adjusted to include Pell Grants awards. Assuming that every individual that satisfies Pell Grants eligibility will obtain it, we subtract Pell Grants awards from tuition cost for each MSA. The Pell Grants awards were calculated for each individual given their characteristics according to the Regular Payment and Disbursement Schedules of Pell Grants from the Department of Education from 1979-2000. These schedules determine Pell Grant dollar amounts for which an individual is eligible given the cost of attendance and an expected family contribution (EFC). The cost of attendance was approximated with the local tuition variable constructed from HEGIS and IPEDS for a given year. The EFC calculation requires a large amount of individual data such as family income, parental assets, family size, number of kids in college, emergency expenses, etc. We approximate EFC using family income in 1979, family size, and the type of student (dependent or independent).

B Estimation Procedure for Control Function Model

This appendix explains the estimation procedure of the control function approach introduced by Heckman (1976, 1980), and Heckman and Robb (1985, 1986). For convenience we outline the complete identification strategy of the model, which relies heavily on Hansen, Heckman, and Mullen (2004). The main difference is that we explicitly account for covariates X, in addition to other differences in the estimation due to the nature of the sample. First, due to reduced sample sizes of blacks and Hispanics, the normalization of the factor loadings $\lambda(S_t)$ is performed in 10th grade at the test date instead of 9th grade or lower. Second, we derive the estimation procedure when the test score depends explicitly on school quality and family characteristics that could affect test scores.

B.1 Identification of the Model and Covariates

The previous calculation implicitly conditions on other covariates that affect test scores at a given date. Achievement test scores could be influenced by environmental characteristics and family backgrounds. In this section, we extend the control function approach to depend on other factors X, other than schooling at test date and final schooling.

Rewrite the test equation (1) from section 3 to depend explicitly on covariates X:

$$T\left(S_T, X\right) = \mu\left(S_T, X\right) + \lambda\left(S_T, X\right) f + \varepsilon\left(S_T\right) \tag{6}$$

where for $\mu\left(S_T, X\right)$ and $\lambda\left(S_T, X\right)$ dependence on X is not necessarily linear. We assume that $\varepsilon \perp\!\!\!\perp X \mid S$. The model assumes that latent ability is independent of environmental factors, $f \perp\!\!\!\perp X$. Test scores conditional on final schooling and schooling at test date and X can be written as:

$$E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] = \mu\left(s_T, X\right) + \lambda\left(s_T, X\right) E\left[f \mid S = s, X\right]$$

where $E\left[\varepsilon\left(S_T\right) \mid S_T = s_T, S = s, X = x\right] = 0$ reflects the assumption that test scores are not used to make schooling decisions. In Hansen, Heckman, and Mullen (2004) and Appendix A for the complete derivation of the identification strategy of the model, individuals consider only final schooling when making schooling decisions, thus $E\left[f \mid S_T = s_T, S = s\right] = E\left[f \mid S = s\right]$. In this case, we present a more stringent assumption $E\left[f \mid S_T = s_T, S = s, X\right] = E\left[f \mid S = s\right]$. The condition $f \perp\!\!\!\perp X$, does not suffice since f conditional on $S = s$ is not necessarily independent of X since $S \not\perp\!\!\!\perp X$. The necessary condition is X to be random with respect to latent ability conditional on final schooling $f \perp\!\!\!\perp X \mid S$.

For two final schooling levels $s, s\prime \geq s_T$

$$
\begin{aligned}
& E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T, S = s', X\right] \\
& = \lambda\left(s_T, X\right)\left[E\left[f \mid S = s, X\right] - E\left[f \mid S = s', X\right]\right] \qquad .
\end{aligned}
$$

Given two values of schooling at test date $s_T, s_T' \leq s, s'$ and $\lambda\left(s_T', x\right) \neq 0$, we can form the ratio

$$
\frac{E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T, S = s', X\right]}{E\left[T\left(S_T\right) \mid S_T = s_T', S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T', S = s', X\right]} = \frac{\lambda\left(s_T, X\right)}{\lambda\left(s_T', X\right)}
$$

with a normalization (in our case $\lambda(2, X) = 1$) we can identify all $\lambda\left(s_T, X\right), s_T = 1, \ldots, \bar{S} - 1$.

With these values we can identify $\mu(s_T, X)$, consider the expected test scores conditional on S_T

$$
E\left[T\left(S_T\right) \mid S_T = s_T, X\right] = \mu\left(s_T, X\right) + \lambda\left(s_T, X\right) E\left[f \mid S_T = s_T\right].
$$

subtracting this from equation (2) we obtain

$$
E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T, X\right] = \lambda\left(s_T, X\right)\left[E\left[f \mid S = s\right] - E\left[f \mid S_T = s_T\right]\right]
$$

Rearranging, we obtain

$$
E\left[f \mid S = s\right] - E\left[f \mid S_T = s_T\right] = \frac{E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right] - E\left[T\left(S_T\right) \mid S_T = s_T, X\right]}{\lambda\left(s_T, X\right)} \qquad (7)
$$

where the RHS can be estimated from the data. Following Hansen, Heckman, and Mullen (2004), we let $E\left[f \mid S = s\right] = a_s$ and $E\left[f \mid S_T = s_T\right] = b_{s_T}$. For different values of S_T and S, we can identify the following combination of parameters assorted in a matrix:

$$\begin{pmatrix} a_{\bar{S}}-b_1 & a_{\bar{S}-1}-b_1 & a_{\bar{S}-2}-b_1 & \ldots & \ldots & a_1-b_1 \\ a_{\bar{S}}-b_2 & a_{\bar{S}-1}-b_2 & \ldots & \ldots & a_2-b_2 & \sim \\ \vdots & \vdots & \ldots & \ldots & \sim & \sim \\ a_{\bar{S}}-b_{\bar{S}-1} & a_{\bar{S}-1}-b_{\bar{S}-1} & \sim & \sim & \sim & \sim \end{pmatrix} \tag{8}$$

where the character "$\sim$" in a cell indicates non-existence of such combination since $s_T \leq s$.

Due to the normalization of latent ability $E(f) = 0$, this can be expressed as two conditions

$$E(f) = E\left[E\left[f \mid S = s\right] \mid S = s\right] = 0 \tag{9}$$

$$= \sum_{j=1}^{\bar{S}} E\left[f \mid S = s\right] \Pr(S = j) = 0 \tag{10}$$

$$E(f) = E\left[E\left[f \mid S_T = s_T\right] \mid S_T = s_T\right] = 0 \tag{11}$$

$$= \sum_{j=1}^{\bar{S}_T} E\left[f \mid S_T = s_T\right] \Pr(S_T = j) = 0. \tag{12}$$

In our context, we define $P_j = \Pr(S = j)$ and express the previous conditions as

$$\sum_{j=1}^{\bar{S}} a_j P_j = 0 \tag{13}$$

and define $\widetilde{P}_j$=$\Pr(S_T = j)$,

$$\sum_{j=1}^{\bar{S}_T} b_j \widetilde{P}_j = 0. \tag{14}$$

Summing across the first row of matrix (8), we obtain

$$\sum_{j=1}^{\bar{S}} P_j(a_j - b_1) = \sum_{j=1}^{\bar{S}} a_j P_j - b_1 \sum_{j=1}^{\bar{S}} P_j = -b_1$$

from equation (13) and the fact that $\sum_{j=1}^{\bar{S}} P_j = 1$. Since all the combinations in the matrix are identified, with b_1 in hand we can identify a_j, for $j = 1, \ldots, \bar{S}$ from the first row. From the identification of a_j, for $j = 1, \ldots, \bar{S}$ we obtain the remaining b_j, $j = 1, \ldots, \bar{S} - 1$ going across each column. Using (14), we identify $b_{\bar{S}}$. Thus, the model is fully identified except for $\lambda\left(\bar{S}, X\right)$. The model is overidentified, for each value b_j we can obtain a combination a_j, for $j = 1, \ldots, \bar{S}$ that satisfy each row. Hence, we use a minimum distance approach to estimate the model with overidentifying restrictions.

Although, the parameter $E\left[T\left(S_T\right) \mid S_T = s_T, S = s, X\right]$ is not necessarily linear, we approximate this parameter by a linear regression for each category of S, S_T and each demographic group.

$$T_{S_T,S} = X\beta_{S_T,S} + \varepsilon_{S_T,S} \tag{15}$$

The expected test scores is given by

$$\hat{T}_{S_T,S} = \bar{X}_{S_T}\hat{\beta}_{S_T,S}$$

where all the persons of a demographic group and S_T have the same level of the environmental variable $\bar{X}_{S_T}$. Although it is not possible to explain variances within each demographic group at a particular level of schooling at the test date, we can analyze the overall effect between demographic groups. Now $a_{ij} = E\left[T\left(S_T\right) \mid S_T = i, S = j, X\right] = \hat{T}_{i,j}$, and from this point we calculate the ratios of the factor loadings and follow the procedure outlined in the previous section keeping the condition on X.

B.2 Empirical Implementation

This section outlines the control function assuming random entry into schooling. The individuals are grouped into six categories of schooling at test date and four categories of final schooling. The categories of schooling at test date are: 9th grade or less ($S_T = 1$), 10th grade, 11th grade, 12th grade, 13 or 14 years of schooling, 15 or more years. Final schooling status are: high-school dropout and GED holder ($S = 1$), high-school graduate, those with some college or an associate degree, and college graduate. We can place each individual in a given cell of a matrix with schooling at test date in the rows and final schooling in the columns:

$$\begin{array}{cc} & S \\ S_T & \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ \sim & a_{42} & a_{43} & a_{44} \\ \sim & \sim & a_{53} & a_{54} \\ \sim & \sim & \sim & a_{64} \end{pmatrix} \end{array} = A$$

where $a_{ij} = E\left[T(i) \mid S_T = i, S = j\right]$ represents the average test score for individuals with i level schooling at test date and level j of final schooling. The character "~" means that no observations are available for that cell [31].

Following the procedure outlined in section 3 we can only identify the ratios of the factor loadings. Due to sample restrictions for minorities, we normalize all the factor loadings relative to the factor that corresponds to 10th grade at test date $\lambda(S_T = 2) = \lambda_2 = 1$. The control function estimation relies on differences in final schooling for those with the same schooling at the test date to identify the factor loadings. The definition of schooling categories is crucial to the number of conditions that we obtain for each level of schooling at test date. We have six conditions that identify the factor for high-school dropouts λ_1:

$$\begin{aligned} \lambda_1 &= \frac{a_{11} - a_{1j}}{a_{21} - a_{2j}}, \quad j = 2, 3, 4 \\ \lambda_1 &= \frac{a_{12} - a_{1j}}{a_{22} - a_{2j}}, \quad j = 3, 4 \\ \lambda_1 &= \frac{a_{13} - a_{1j}}{a_{23} - a_{2j}}, \quad j = 4 \end{aligned}$$

Let $Y_1(A)$ be the vector of the six unrestricted estimators $Y_1(A) = [\lambda_1(1) \ldots \lambda_1(6)]'$. Since we have a set of six conditions that identify the factor λ_1, we obtain the optimal λ_1 with a minimum distance

[31] The cell "41" represented GED recipients that are coded as 12th graders, but we categorize them as high-school dropouts. The estimation procedure does not use this category; therefore, its code is not available.

framework, which is the one that minimizes:

$$q(\lambda_1) = (Y_1(A) - \iota\lambda_1)' W_1 (Y_1(A) - \iota\lambda_1)$$

where W_1 is the inverse of an estimate of the asymptotic covariance matrix of $Y_1(A)$[32]. The solution is represented by

$$\hat{\lambda}_1 = (\iota' W_1 \iota)^{-1} \iota' W_1 Y_1(A).$$

For individuals with 11 years of schooling, we have identifying conditions similar to those used for individuals in 9th grade or lower. In this case the conditions for $\lambda_3$are:

$$\lambda_3 = \frac{a_{31} - a_{3j}}{a_{21} - a_{2j}}, \quad j = 2,3,4$$
$$\lambda_3 = \frac{a_{32} - a_{3j}}{a_{22} - a_{2j}}, \quad j = 3,4$$
$$\lambda_3 = \frac{a_{33} - a_{3j}}{a_{23} - a_{2j}}, \quad j = 4$$

where the optimal λ_3 is calculated using the same procedure as λ_1.

[32]The inverse of the asymptotic matrix is constructed as follows:

$$W_1 = (\nabla_{\lambda_1 a}\ V_1\ \nabla'_{\lambda_1 a})^{-1}$$

Where $\nabla_{\lambda_1 a}$ is the gradient matrix of the estimator:

$$\nabla_{\lambda_1 a} = \begin{pmatrix} \frac{\partial\lambda_1(1)}{\partial a_{21}} & \frac{\partial\lambda_1(1)}{\partial a_{22}} & \frac{\partial\lambda_1(1)}{\partial a_{23}} & \frac{\partial\lambda_1(1)}{\partial a_{24}} & \frac{\partial\lambda_1(1)}{\partial a_{11}} & \frac{\partial\lambda_1(1)}{\partial a_{11}} & \frac{\partial\lambda_1(1)}{\partial a_{13}} & \frac{\partial\lambda_1(1)}{\partial a_{14}} \\ \frac{\partial\lambda_1(2)}{\partial a_{21}} & \frac{\partial\lambda_1(2)}{\partial a_{22}} & & & & & & \vdots \\ \vdots & & \ddots & & & & & \\ \frac{\partial\lambda_1(6)}{\partial a_{21}} & \cdots & & & & & & \frac{\partial\lambda_1(6)}{\partial a_{14}} \end{pmatrix}$$

and the sample variance-covariance matrix:

$$V_1 = \begin{pmatrix} \sigma_{21}^2 & 0 & \cdots & & 0 \\ 0 & \sigma_{22}^2 & & & \vdots \\ \vdots & & \ddots & & \\ & & & \sigma_{13}^2 & 0 \\ 0 & & & 0 & \sigma_{14}^2 \end{pmatrix}$$

where σ_{ij}^2 are the sample variance of the group means a_{ij}

$$\sigma_{ij}^2 = \frac{\sum_{k=1}^{n_{ij}} (t_{kij} - a_{ij})^2}{n_{ij}(n_{ij} - 1)}$$

where t_{kij} are the test scores of the individual k in the group (i,j).

High school graduates (12 years of school completed) at the test date can only be observed at three levels of final education (high-school, some college and college graduates). In this case the conditions to identify the loadings of some college at test date $\lambda_4$are:

$$\lambda_4 = \frac{a_{42} - a_{43}}{a_{22} - a_{23}}$$
$$\lambda_4 = \frac{a_{42} - a_{44}}{a_{22} - a_{24}}$$
$$\lambda_4 = \frac{a_{43} - a_{44}}{a_{23} - a_{24}}.$$

For individuals with some college or associate degree at the time of the test, we only observe some college or college graduates as final schooling. Hence, we have only one condition for λ_5:

$$\lambda_5 = \frac{a_{53} - a_{54}}{a_{23} - a_{24}}.$$

The factor loading for those college graduates, λ_6, is not identified.

The next step is to estimate the intercept of the test equation, $\mu(S_T)$, for each level of schooling at test date and the control function estimates $c_1(s) = E[f \mid S = s]$ for each level of final schooling $s = 1, \ldots, 4$. and $c_2(s_T) = E[f \mid S_T = s_T]$ for each level of schooling at test date $s_T = 1, \ldots, 6$. The model has overidentifying restrictions of expected test scores:

$$E[T(S_T) \mid S_T = 1, S = i] = a_{1i} = \mu(1) + \lambda(1) c_1(i), \; i = 1, 2, 3, 4$$
$$E[T(S_T) \mid S_T = 2, S = i] = a_{2i} = \mu(2) + \lambda(2) c_1(i), \; i = 1, 2, 3, 4$$
$$E[T(S_T) \mid S_T = 3, S = i] = a_{3i} = \mu(3) + \lambda(3) c_1(i), \; i = 1, 2, 3, 4$$
$$E[T(S_T) \mid S_T = 4, S = i] = a_{4i} = \mu(4) + \lambda(4) c_1(i), \; i = 2, 3, 4$$
$$E[T(S_T) \mid S_T = 5, S = i] = a_{5i} = \mu(5) + \lambda(5) c_1(i), \; i = 3, 4$$

and

$$E[T(S_T) \mid S_T = i] = \frac{\sum_{j=1}^{4} n_{ij} a_{ij}}{\sum_{j=1}^{4} n_{ij}} = \mu(i) + \lambda(i) c_2(i), \; i = 1, \ldots, 5$$

where n_{ij} is the sample size of individuals with schooling at test date i and final schooling j. The restrictions that correspond to $E(f) = 0$ are implemented in the following way:

$$E[f \mid S] = \frac{\sum_{j=1}^{4} n_j c_1(j)}{n} = 0 \text{ and } E[f \mid S_T] = \frac{\sum_{j=1}^{6} \tilde{n}_j c_2(j)}{n} = 0$$

where $n_j = \sum_{i=1}^{6} n_{ij}$ is the number of individuals with final schooling j and $\tilde{n}_j$ is the number of individuals with $S_T = j$.

The complete model implies 24 restrictions in the model on 15 parameters

$$\theta = [\mu(1), \ldots \mu(5); c_1(1), \ldots, c_1(4); c_2(1), \ldots, c_2(6)]'$$

which the estimation of these minimizes

$$q(\theta) = (Y(A) - H\theta')' W (Y(A) - \theta)$$

where W is the inverse of an estimate of the covariance matrix of $Y(A)$. Where the optimal GMM will be given by:

$$\hat{\theta} = (H'WH)^{-1} H'WY(A).$$

C Estimation Procedure for Structural Model

Here we briefly describe the main components of the estimation,; please refer to Hansen, Heckman, and Mullen (2004) for a complete derivation of the methodology. The model was estimated via Bayesian Markov Chain Monte Carlo (MCMC) method. Let $S \in \{1, ..., \bar{S}\}$ be the set of final schooling levels, where $S = 1$ represents high school dropout and GEDs, $S = 2$ represents high school graduates, $S = 3$ represents some college and associate degrees, and finally $\bar{S} = 4$ represents college graduates. $Z(s)$ is the set of variables in the s^{th} choice equation, and f is the latent ability factor.

Choice Equations. By independence, the conditional posterior distribution of latent utilities V is the product of the individual conditional posterior distributions of V_i:

$$V_i = \begin{cases} TN_{[\max_{l \neq S}\{V_i(l)\},\infty)}(z_i(s)\gamma(s) + \alpha(s)f_i, 1) & \text{if } S = S_i \\ TN_{(-\infty, V_i(s_i)]}(z(s)\gamma(s) + \alpha(s)f, 1) & \text{if } S \neq S_i \end{cases}$$

where TN represents a truncated normal distribution, the element $\gamma(s)$ follows from a classical linear regression model with non-informative prior:

$$\gamma(s) \sim N(\widehat{\gamma}(s), \widehat{\Omega}(s))$$

where

$$\begin{aligned} \widehat{\gamma}(s) &= (Z(s)'Z(s))^{-1}Z(s)'(V(s) - \alpha(s)f) \\ \widehat{\Omega}(s) &= (Z(s)'Z(s))^{-1}. \end{aligned}$$

We assume a standard normal prior distribution for the parameters $\{\alpha(s)\}_{s=1}^{\overline{S}-1} \sim N(0, 1)$ for each s. To assure identification, we set $\alpha(s = 2) = 0$ and $\gamma(s = 2) = 0$. The conditional distributions of $\alpha(s)$ is then:

$$\alpha(s) \sim N\left(\widehat{\alpha}(s), \widehat{\Omega_1}(s)\right)$$

where

$$\widehat{\alpha}(s) = (f'f+1)^{-1}\left(f'\left(V(s) - Z(s)\gamma\left(s\right)\right)\right)$$
$$\widehat{\Omega_1}(s) = (f'f+1)^{-1}.$$

The vector $Z(s)$ includes the following variables: lived in an urban setting at age 14, broken home status at age 14, number of siblings in 1979, lived in the south at age 14, mother's education with imputed values, father's education with imputed values, family income at year 1979 with imputed values and year of birth dummies. Finally, local wage and local unemployment rate by level of schooling. In the case of the normalize level, $s = 2$ (high school graduates), the vector $Z(s)$ includes only the local wage and the local unemployment rate.

Test Equations. The test equations

$$T_k^*(S_T) = X(S_T)\beta_k(S_T) + \lambda_k(S_T)f + \varepsilon_k(S_T) \qquad k = 1, ..., 4$$

corresponds to the standardized ASVAB tests: word knowledge, paragraph comprehension, math knowledge and arithmetic reasoning. For each schooling level at the time of the test $S_T = 1, ..., \bar{S}_T$, we estimate the coefficients:

$$\beta(s) \sim N(\widehat{\beta}(s), \widehat{\Omega_2}(s)).$$

where

$$\widehat{\beta}(s) = (X(s)'X(s))^{-1}X(s)'(T_k^*(s) - \lambda_k(s)f)$$
$$\widehat{\Omega_2}(s) = \sigma_k^2(s)(X(s)'X(s))^{-1}$$

Since we can estimate $\lambda_k(s)f$ we improve from the case were f is unobservable. We assume a prior distribution of $\{\lambda_k(S)\}_{s=1,k=1}^{\overline{S}_T,K} \sim N(0,1)$. In addition, to assure identification, we set $\widehat{\lambda_1}\left(s_T = 1\right) = 1$.

The conditional distribution is given by

$$\lambda_k(S) \sim N(\widehat{\lambda_k}(s), \widehat{\Omega_k}(s))$$

where

$$\begin{aligned}
\widehat{\lambda_k}(s) &= \widehat{\Omega_k}(s) \left[\frac{f' \left(T_k^*(s) - X(S_T)\beta_k(s_T) \right)}{\sigma_k^2(s)} \right] \\
\widehat{\Omega_k}(s) &= \left[\frac{f'f}{\sigma_k^2(s)} + 1 \right]^{-1}.
\end{aligned}$$

Define the estimate of $\varepsilon_k(S_T)$ by

$$e_{k(s)} = T_k^*(s_T) - X(s_T)\beta_k(s_T) - \lambda_k(s_T)f$$

Let $\sigma_k^2(s)$ be the variance of $\varepsilon_k(S_T)$, we assume its prior as an Inverse Gamma (IG) distribution

$$\sigma_k^2(s) \sim IG\left(\frac{n(s)}{2} + a_s, \frac{e'_{k(s)}e_{k(s)}}{2} + b_s \right)$$

where we set $a_s = 2$ and $b_s = 1$.

We account for ceiling effects on test scores. Individuals who reach the maximum score of a test are sampled from the following truncated normal:

$$T_{ik}^* \sim TN_{[c_k,\infty)} \left(x_i'(s_{T_i})\beta_k(s_{T_i}) + \lambda_k(s_{T_i})f_i, \sigma_k^2(s_{T_i}) \right)$$

Factors. The factor f is assumed to be derived from a mixture of normals:

$$f \sim \sum_{i=1}^{I} p_i N(\mu_{g_i}, \sigma_{g_i}^2)$$

where $g_i \in \{1, ..., I\}$ denote the mixture components from which f_i is sampled. Conditional on g_i the conditional distribution of f_i is:

$$f_i \sim N(\widehat{f}_i, \widehat{\sigma}_i).$$

where

$$\widehat{f}_i = \widehat{\sigma}_i \left[\sum_{s=1}^{\overline{S}} \alpha^2(s) + \sum_{k=1}^{K} \frac{\lambda_k(s_{T_i})}{\sigma_k^2(s_{T_i})} \left(T_{ik}^*(s) - X_i(S_{T_i})\beta_k(S_{T_i})\right) + \frac{1}{\sigma_{g_i}^2}\mu_{g_i} \right]$$

$$\widehat{\sigma}_i = \left[\sum_{s=1}^{\overline{S}} \alpha^2(s) + \sum_{k=1}^{K} \frac{\lambda_k(s_{T_i})}{\sigma_k^2(s_{T_i})} + \frac{1}{\sigma_{g_i}^2} \right]^{-1}$$

We impose the restriction $\sum_{i=1}^{I} p_i \mu_i = 0$. We further assume $I = 3$.

References

Betts, J. R. (1995, May). Does school quality matter? Evidence from the National Longitudinal Survey of Youth. *Review of Economics and Statistics 77*(2), 231–250.

Boozer, M. A., A. B. Krueger, and S. Wolkon (1992). Race and school quality since Brown v. Board of Education. *Brookings Papers on Economic Activity*, 269–326.

Carneiro, P., J. J. Heckman, and D. V. Masterov (2005, April). Labor market discrimination and racial differences in pre-market factors. *Journal of Law and Economics 47*(1), 1–39.

Ceci, S. J. (1991, September). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology 27*(5), 703–722.

Duncan, G. J. and J. Brooks-Gunn (1997). *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.

Ferguson, R. (2002a). What doesn't meet the eye: Understanding and addressing racial disparities in high acheiving suburban schools. Special Edition, Policy Issues Report (Naperville, IL: North Central Regional Educational Laboratory).

Ferguson, R. (2002b). Why America's black-white school acheivement gap persists. Unpublished manuscript, Harvard University.

Fryer, R. and S. Levitt (2004, May). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics 86*(2), 447–464.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004, July-August). The effect of schooling and ability on achievement test scores. *Journal of Econometrics 121*(1-2), 39–98.

Hanushek, E. A. (1989, May). Expenditures, efficiency, and equity in education: The federal government's role. *American Economic Review 79*(2), 46–51.

Hanushek, E. A. (2004, Special Issue, May). What if there are no 'best practices'? *Scottish Journal of Political Economy 51*(2), 156–172.

Heckman, J. J. (1976). Simultaneous equation models with both continuous and discrete endogenous variables with and without structural shift in the equations. In S. Goldfeld and R. Quandt (Eds.), *Studies in Nonlinear Estimation*, pp. 235–272. Cambridge, MA: Ballinger Publishing Company.

Heckman, J. J. (1980). Addendum to sample selection bias as a specification error. In E. Stromsdorfer and G. Farkas (Eds.), *Evaluation Studies Review Annual*, Volume 5. Beverly Hills: Sage Publications.

Heckman, J. J., A. Layne-Farrar, and P. E. Todd (1996, July). Does measured school quality really matter? An examination of the earnings-quality relationship. In G. Burtless (Ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Brookings Institution Press.

Heckman, J. J. and D. Neal (1996, February). Coleman's contributions to education: Theory, research styles and empirical research. In J. Clark (Ed.), *James S. Coleman*, Consensus and Controversy Falmer Sociology Series, pp. 81–102. Falmer Press.

Heckman, J. J. and R. Robb (1985, October-November). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics 30*(1-2), 239–267.

Heckman, J. J. and R. Robb (1986). Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes. In H. Wainer (Ed.), *Drawing Inferences from Self-Selected Samples*, pp. 63–107. New York: Springer-Verlag. Reprinted in 2000, Mahwah, NJ: Lawrence Erlbaum Associates.

Herrnstein, R. J. and C. A. Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

Neal, D. A. and W. R. Johnson (1996, October). The role of premarket factors in black-white wage differences. *Journal of Political Economy 104*(5), 869–895.

Phillips, M., J. Brooks-Gunn, G. Duncan, P. K. Klebanov, and J. Crane (1998). Family background, parenting practices and the black-white test score gap. In C. Jenks and M. Phillips (Eds.), *The Black-White Test Score Gap*, pp. 103–145. Washington, DC: Brookings Institution Press.

Phillips, M., J. Crouse, and J. Ralph (1998). Does the black-white test score gap widen after children enter school? In C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap*, pp. 401–427. Washington, DC: The Brookings Institution.

Winship, C. and S. Korenman (1997). Does staying in school make you smarter? The effect of education on IQ in *The Bell Curve*. In B. Devlin, S. Feinberg, D. Resnick, and K. Roeder (Eds.), *Intelligence, Genes, and Success: Scientists Respond to The Bell Curve*, pp. 215–234. New York: Springer, Copernicus.

Table 1

Educational Variables by Race and Gender, NLSY79

	White		Black		Hispanic	
	Male	Female	Male	Female	Male	Female
Number of Observations	2063	2137	1192	1235	717	774
Panel A. Schooling Level at Test Date						
Less than or Equal to 9 Years of Schooling	0.09	0.06	0.15	0.10	0.20	0.17
	(0.29)	(0.25)	(0.36)	(0.30)	(0.40)	(0.38)
10th Grade	0.16	0.15	0.19	0.16	0.19	0.18
	(0.37)	(0.35)	(0.39)	(0.37)	(0.39)	(0.38)
11th Grade	0.15	0.13	0.18	0.15	0.18	0.16
	(0.35)	(0.34)	(0.39)	(0.36)	(0.39)	(0.37)
High School	0.39	0.41	0.34	0.37	0.30	0.33
	(0.49)	(0.49)	(0.48)	(0.48)	(0.46)	(0.47)
13-14 Years	0.14	0.18	0.10	0.18	0.11	0.12
	(0.35)	(0.38)	(0.30)	(0.38)	(0.31)	(0.33)
15 or More Years	0.07	0.07	0.03	0.04	0.03	0.03
	(0.26)	(0.26)	(0.17)	(0.19)	(0.16)	(0.18)
Panel B. Age and Enrollment						
Enrolled Status at Test Date	0.51	0.44	0.45	0.46	0.46	0.42
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)
Age at December 31st, 1980	19.16	19.39	19.05	19.27	18.95	19.11
	(2.17)	(2.18)	(2.12)	(2.17)	(2.14)	(2.15)
Panel C. Final Schooling Level						
Dropout	0.17	0.13	0.29	0.19	0.33	0.25
	(0.37)	(0.34)	(0.45)	(0.39)	(0.47)	(0.43)
High School	0.35	0.35	0.38	0.33	0.32	0.30
	(0.48)	(0.48)	(0.49)	(0.47)	(0.47)	(0.46)
Some College +2 Year College	0.20	0.24	0.21	0.33	0.24	0.31
	(0.40)	(0.43)	(0.41)	(0.47)	(0.43)	(0.46)
4 Year of College or More	0.29	0.28	0.12	0.15	0.11	0.13
	(0.45)	(0.45)	(0.33)	(0.36)	(0.32)	(0.34)
Panel D. AFQT and ASVAB Components						
Arithmetic Reasoning	19.63	17.86	11.69	10.93	14.33	12.58
	(7.19)	(6.74)	(5.44)	(4.68)	(6.61)	(5.46)
Word Knowledge	26.78	27.07	17.43	18.32	21.05	20.58
	(7.06)	(6.57)	(8.32)	(7.60)	(8.23)	(7.62)
Paragraph Comprehension	10.94	11.71	7.41	8.24	8.58	9.14
	(3.30)	(2.83)	(3.65)	(3.45)	(3.70)	(3.46)
Math Knowledge	14.55	14.02	9.04	9.23	10.53	9.85
	(6.52)	(6.05)	(4.85)	(4.58)	(5.79)	(5.30)
AFQT	71.90	70.66	45.65	46.72	54.63	52.21
	(21.55)	(19.46)	(19.41)	(17.54)	(21.52)	(18.90)

Notes: Standard deviations in parentheses. AFQT is the sum of arithmetic reasoning, word knowledge, paragraph comprehension and math knowledge ASVAB scores.

Table 2

Age-Adjusted AFQT Distributions Fitted in the White Quartile Distribution

	Q1	Q2	Q3	Q4
Panel A. All Individuals				
White	25	25	25	25
Black	73	17	7	3
Hispanic	59	23	11	7
Panel B. Less than or Equal to 9 Years				
White	25	25	25	25
Black	55	30	11	4
Hispanic	39	36	15	10
Panel C. 10th Grade				
White	25	25	25	25
Black	70	22	6	2
Hispanic	55	27	12	5
Panel D. 11th Grade				
White	26	25	24	24
Black	74	17	7	2
Hispanic	55	24	12	9
Panel E. High School				
White	25	25	25	25
Black	76	14	6	4
Hispanic	57	21	13	9
Panel F. 13-14 Years				
White	25	25	25	25
Black	82	11	4	3
Hispanic	66	17	12	5
Panel G. 15 Years or More				
White	25	25	25	25
Black	73	18	6	4
Hispanic	51	22	16	11

Note: Schooling at test date. The quartiles distribution was based in the white distribution. The quartiles of the white AFQT distribution for each school level define the intervals where the blacks and minorities's AFQT are compared. The percentage represents the proportion of individuals whose AFQT scores lie in the given quartile.

Table 3

Percentage of Schooling Attainment at Test Date by Age, NLSY79

	Highest Grade Completed at Test Date, S_T					
Age at Test Date	Less than or Equal to 9	10th Grade	11th Grade	High School	13-14 Years	15 Years or More
Panel A. White Males						
16	32.83	66.04	0.75	0.38	0.00	0.00
17	8.22	30.48	55.14	5.82	0.34	0.00
18	8.45	8.16	23.03	59.48	0.87	0.00
19	5.07	3.26	7.25	57.97	26.45	0.00
20	3.87	4.58	5.28	53.17	32.04	1.06
21	5.22	3.48	3.91	49.57	22.17	15.65
22	3.21	2.67	4.28	41.18	19.25	29.41
23	5.91	3.23	5.38	39.25	18.28	27.96
Panel B. White Females						
16	17.83	80.43	0.87	0.87	0.00	0.00
17	7.67	21.60	63.41	6.97	0.35	0.00
18	5.33	6.33	17.33	70.00	1.00	0.00
19	4.73	5.36	4.10	54.89	30.91	0.00
20	5.57	3.14	4.53	43.55	39.37	3.83
21	1.88	2.63	3.76	45.49	27.07	19.17
22	4.39	3.95	0.88	52.63	18.86	19.30
23	5.86	1.80	3.15	46.40	20.27	22.52
Panel C. Black Males						
16	48.70	48.70	2.60	0.00	0.00	0.00
17	17.39	35.87	41.85	4.89	0.00	0.00
18	10.36	10.88	31.61	46.11	1.04	0.00
19	8.15	9.24	17.39	48.37	16.85	0.00
20	7.55	8.81	9.43	55.35	18.87	0.00
21	8.94	12.20	8.94	42.28	20.33	7.32
22	7.21	13.51	8.11	40.54	17.12	13.51
23	6.98	3.49	12.79	45.35	16.28	15.12
Panel D. Black Females						
16	28.38	68.24	2.70	0.68	0.00	0.00
17	10.78	28.14	56.89	4.19	0.00	0.00
18	7.69	7.69	20.33	62.09	2.20	0.00
19	8.79	3.30	9.34	51.10	27.47	0.00
20	6.90	5.17	8.05	46.55	31.03	2.30
21	6.76	8.11	3.38	43.24	29.73	8.78
22	4.27	7.69	5.13	35.04	31.62	16.24
23	5.13	4.27	6.84	46.15	27.35	10.26
Panel E. Hispanic Males						
16	50.50	47.52	0.99	0.99	0.00	0.00
17	26.72	31.30	38.93	3.05	0.00	0.00
18	16.00	18.00	29.00	37.00	0.00	0.00
19	10.19	11.11	17.59	46.30	14.81	0.00
20	8.42	6.32	12.63	53.68	17.89	1.05
21	8.00	9.33	16.00	40.00	24.00	2.67
22	9.26	5.56	7.41	38.89	27.78	11.11
23	18.18	3.64	3.64	34.55	20.00	20.00
Panel F. Hispanic Females						
16	42.70	56.18	1.12	0.00	0.00	0.00
17	17.56	35.11	43.51	3.82	0.00	0.00
18	9.09	12.12	25.76	49.24	3.79	0.00
19	9.71	7.77	7.77	58.25	16.50	0.00
20	20.83	6.25	13.54	37.50	21.88	0.00
21	13.79	4.60	8.05	45.98	22.99	4.60
22	11.27	5.63	4.23	39.44	28.17	11.27
23	16.67	4.55	3.03	34.85	19.70	21.21

Table 4

Sample Means of Background Characteristics, NLSY79

	White		Black		Hispanic	
	Male	Female	Male	Female	Male	Female
Number of observations	2063	2137	1194	1235	719	775
Lived in Urban Age 14	0.75	0.76	0.84	0.82	0.89	0.90
	(0.44)	(0.43)	(0.37)	(0.39)	(0.31)	(0.30)
Broken Home Status at Age 14	0.20	0.22	0.50	0.51	0.33	0.32
	(0.40)	(0.41)	(0.50)	(0.50)	(0.47)	(0.47)
Number of Siblings 79	2.99	3.09	4.73	4.75	4.43	4.49
	(1.99)	(1.95)	(2.99)	(3.02)	(2.81)	(3.00)
Lived in South Age 14	0.25	0.27	0.54	0.58	0.27	0.28
	(0.43)	(0.45)	(0.50)	(0.49)	(0.44)	(0.45)
Mother's Education	12.04	11.92	10.86	10.68	8.06	8.06
	(2.31)	(2.36)	(2.38)	(2.62)	(4.07)	(3.87)
Father's Education	12.35	12.19	10.24	10.17	8.49	8.21
	(3.21)	(3.09)	(2.93)	(3.22)	(4.26)	(4.32)
Family Income in 1979 in 1,000	22.30	21.55	11.70	11.10	13.98	13.11
	(13.13)	(12.40)	(8.21)	(8.89)	(10.02)	(8.73)
Age at December 31st, 1980	19.16	19.39	19.05	19.27	18.95	19.11
	(2.17)	(2.18)	(2.12)	(2.17)	(2.14)	(2.15)

Note: Standard deviations in parentheses.

Table 5
Sample Means of School Quality Measures in 1979, NLSY79

	White		Black		Hispanic	
	Male	Female	Male	Female	Male	Female
Books per Student	15.42	14.14	11.84	11.93	11.70	13.45
	(35.89)	(8.80)	(10.04)	(7.55)	(9.49)	(28.20)
	1482	1431	639	653	415	455
Teacher Turnover Excluding Death or Retirement	6.42	6.68	6.71	6.80	8.44	9.39
	(8.02)	(8.48)	(7.82)	(7.87)	(10.24)	(11.04)
	1653	1633	762	794	465	498
Annual Salary for a New Certified Teacher With a BA in $100, 2000 dollars	108.19	107.86	109.36	108.44	109.06	106.88
	(10.82)	(11.21)	(14.78)	(15.25)	(10.56)	(11.89)
	1603	1575	747	761	455	486
Average Daily Percentage Attendance in the School	89.90	89.38	85.19	85.24	86.60	87.90
	(14.58)	(15.40)	(17.67)	(17.45)	(14.89)	(12.87)
	1656	1635	764	803	484	501
Percentage of Faculty with MAs or PhDs	47.81	48.09	49.77	50.10	46.56	46.47
	(22.92)	(22.90)	(23.31)	(23.38)	(24.56)	(23.05)
	1630	1605	769	796	467	483
Full-Time Equivalent Teachers Per 100 Students	5.98	5.52	5.47	5.44	5.22	5.13
	(15.02)	(1.53)	(3.63)	(2.53)	(1.59)	(1.57)
	1582	1545	762	792	461	494
Full-Time Equivalent Counselors Per 100 Students	0.36	0.33	0.33	0.33	0.31	0.31
	(1.07)	(0.53)	(0.51)	(0.51)	(0.20)	(0.27)
	1604	1574	769	812	461	486

Note: Standard deviations in parentheses.The number of observations is displayed below the standard deviation.

Table 6

Nonparametric Estimates of $\mu(S_T)$ of Control Funtion Model

	Males			Females		
	White	Black	Hispanic	White	Black	Hispanic
Panel A. Test Scores Intercepts						
$\mu(S_T$ =9th Grade or Less)	52.85	31.95	39.28	53.61	33.61	44.84
	(1.20)	(0.80)	(1.12)	(1.41)	(0.87)	(1.29)
$\mu(S_T$=10th Grade)	66.14	42.40	49.97	66.60	44.81	49.58
	(0.94)	(0.93)	(1.31)	(0.88)	(0.91)	(1.37)
$\mu(S_T$=11th Grade)	71.08	45.74	56.22	68.52	44.04	53.06
	(0.82)	(0.98)	(1.51)	(0.94)	(0.95)	(1.35)
$\mu(S_T$=High School)	74.39	46.84	59.60	71.39	48.25	53.70
	(0.64)	(0.99)	(1.43)	(0.55)	(0.76)	(1.07)
$\mu(S_T$=Some College, 13-14 Years)	79.31	50.49	71.92	78.54	45.50	54.22
	(0.77)	(1.88)	(2.05)	(0.70)	(1.31)	(1.94)
Panel B. Causal Effect of Schooling $\mu(S_T)$-$\mu(S'_T)$						
$\mu(S_T$=10th Grade)- $\mu(S_T$ =9th Grade or lower)	13.29	10.45	10.69	12.99	11.20	4.74
$\mu(S_T$=11th Grade)-$\mu(S_T$ =10th Grade)	4.94	3.34	6.26	1.92	-0.77	3.48
$\mu(S_T$=High School)- $\mu(S_T$ =11th Grade)	3.31	1.10	3.38	2.87	4.21	0.64
$\mu(S_T$=Some College, 13-14 Years)- $\mu(S_T$ =High School)	4.92	3.65	12.32	7.15	-2.75	0.52

Note: Standard errors in parentheses.

Table 7
Nonparametric Estimates of Factor Loadings

	Comparison	Males			Females		
	Group (s,s')	White	Black	Hispanic	White	Black	Hispanic
λ (S_T=9th Grade or Less)	(1,2)	0.38	-0.21	1.37	-0.22	0.05	2.29
λ (S_T=9th Grade or Less)	(1,3)	0.62	0.78	0.28	1.20	0.44	0.70
λ (S_T=9th Grade or Less)	(1,4)	0.95	0.38	0.41	0.91	0.37	1.47
λ (S_T=9th Grade or Less)	(2,3)	0.79	1.22	-0.15	2.55	0.61	0.05
λ (S_T=9th Grade or Less)	(2,4)	1.14	0.51	0.25	1.29	0.45	1.29
λ (S_T=9th Grade or Less)	(3,4)	1.45	-0.14	0.55	0.61	0.19	2.71
λ (S_T=9th Grade or Less)$_{MD}$		0.72	0.15	0.50	0.83	0.37	1.35
		(0.11)	(0.35)	(0.20)	(0.22)	(0.13)	(0.25)
λ (S_T=11th Grade)	(1,2)	0.96	1.80	2.08	1.00	0.73	3.06
λ (S_T=11th Grade)	(1,3)	0.76	1.70	0.92	1.09	0.76	1.53
λ (S_T=11th Grade)	(1,4)	1.00	1.34	1.37	0.95	0.90	1.46
λ (S_T=11th Grade)	(2,3)	0.62	1.66	0.47	1.18	0.78	0.90
λ (S_T=11th Grade)	(2,4)	1.01	1.24	1.25	0.93	0.94	1.10
λ (S_T=11th Grade)	(3,4)	1.37	0.87	1.85	0.80	1.19	1.34
λ (S_T=11th Grade)$_{MD}$		0.99	1.41	1.17	0.95	0.83	1.23
		(0.09)	(0.26)	(0.23)	(0.11)	(0.14)	(0.34)
λ (S_T=High School)	(2,3)	0.77	1.43	0.98	0.95	0.83	0.78
λ (S_T=High School)	(2,4)	0.74	1.33	0.93	0.83	0.99	1.01
λ (S_T=High School)	(3,4)	0.71	1.23	0.89	0.77	1.24	1.27
λ (S_T=High School)$_{MD}$		0.74	1.37	0.95	0.83	0.88	0.87
		(0.09)	(0.35)	(0.26)	(0.10)	(0.17)	(0.31)
λ (S_T=13-14 years)	(3,4)	0.80	1.31	0.23	0.59	1.34	1.60
		(0.21)	(0.77)	(0.33)	(0.13)	(0.77)	(1.22)

Note. MD= Minimum distance estimator. We normalize λ (S_T=10th Grade)=1. Standard errors in parentheses.

Table 8

Nonparametric Estimates of Control Functions

	Males			Females		
	White	Black	Hispanic	White	Black	Hispanic
Panel A. Expected Ability Conditional on Schooling at Test Date.						
$E(f \mid S_T$=9th Grade or Less)	-10.77	-2.18	-5.36	-10.80	-8.18	-6.59
	(1.96)	(5.39)	(2.52)	(2.14)	(2.55)	(1.11)
$E(f \mid S_T$=10th Grade)	-2.34	-2.38	-1.68	-1.98	-3.06	-1.42
	(1.29)	(1.21)	(1.78)	(1.19)	(1.18)	(1.90)
$E(f \mid S_T$=11th Grade)	-0.75	-0.48	-0.59	-1.37	-1.63	0.79
	(1.14)	(0.95)	(1.81)	(1.34)	(1.53)	(1.55)
$E(f \mid S_T$=High School)	-2.79	0.15	-0.20	-3.32	-1.19	0.92
	(0.74)	(0.76)	(1.85)	(0.44)	(0.89)	(1.48)
$E(f \mid S_T$=Some College, 13-14 Years)	11.44	9.26	10.16	8.99	9.14	6.64
	(1.29)	(1.92)	(12.73)	(1.57)	(1.22)	(1.55)
Panel B. Expected Ability Conditional on Final Schooling						
$E(f \mid S$=Dropout)	-16.04	-7.44	-8.54	-14.69	-11.64	-8.68
	(2.07)	(1.11)	(1.63)	(2.55)	(1.96)	(1.62)
$E(f \mid S$=High School)	-8.40	-2.85	-3.19	-8.19	-6.68	-2.71
	(0.67)	(0.57)	(1.13)	(0.60)	(0.70)	(0.83)
$E(f \mid S$=Some College)	3.00	6.47	5.36	1.58	6.18	4.04
	(0.85)	(0.74)	(1.34)	(0.72)	(0.69)	(0.83)
$E(f \mid S$=College)	17.55	14.92	22.57	15.72	15.20	13.18
	(0.67)	(1.36)	(2.18)	(0.69)	(1.14)	(1.44)

Note. Standard errors in parentheses.

Table 9

Estimated Gap in AFQT Scores and Its Components by Schooling at Test Date

	Less than or Equal to 9	10th Grade	11th Grade	High School	13-14 Years
Panel A. White Males versus Black Males					
$\mu_W(S_T)-\mu_B(S_T)$	20.90	23.74	25.34	27.56	28.83
$\lambda_W(S_T)E_W(f\mid S_T=s_T) - \lambda_B(S_T)E_B(f\mid S_T=s_T)$	-7.38	0.04	-0.07	-2.27	-2.99
Total Gap	13.52	23.78	25.27	25.28	25.83
Panel B. White Females versus Black Females					
$\mu_W(S_T)-\mu_B(S_T)$	20.00	21.79	24.48	23.13	33.04
$\lambda_W(S_T)E_W(f\mid S_T=s_T) - \lambda_B(S_T)E_B(f\mid S_T=s_T)$	-5.93	1.08	0.06	-1.71	-6.93
Total Gap	14.07	22.87	24.53	21.42	26.11
Panel C. White Males Versus Hispanic Males					
$\mu_W(S_T)-\mu_H(S_T)$	13.58	16.17	14.86	14.79	7.40
$\lambda_W(S_T)E_W(f\mid S_T=s_T) - \lambda_H(S_T)E_H(f\mid S_T=s_T)$	-5.01	-0.66	-0.05	-1.88	6.85
Total Gap	8.56	15.51	14.81	12.91	14.25
Panel D. White Females Versus Hispanic Females					
$\mu_W(S_T)-\mu_H(S_T)$	8.77	17.02	15.46	17.69	24.32
$\lambda_W(S_T)E_W(f\mid S_T=s_T) - \lambda_H(S_T)E_H(f\mid S_T=s_T)$	-0.07	-0.56	-2.27	-3.56	-5.28
Total Gap	8.70	16.45	13.18	14.13	19.04
Panel E. Black Males Versus Hispanic Males					
$\mu_B(S_T)-\mu_H(S_T)$	-7.32	-7.57	-10.48	-12.76	-21.43
$\lambda_B(S_T)E_B(f\mid S_T=s_T) - \lambda_H(S_T)E_H(f\mid S_T=s_T)$	2.37	-0.70	0.02	0.39	9.84
Total Gap	-4.95	-8.26	-10.46	-12.38	-11.59
Panel F. Black Females Versus Hispanic Females					
$\mu_B(S_T)-\mu_H(S_T)$	-11.23	-4.77	-9.02	-5.45	-8.72
$\lambda_B(S_T)E_B(f\mid S_T=s_T) - \lambda_H(S_T)E_H(f\mid S_T=s_T)$	5.85	-1.65	-2.33	-1.84	1.65
Total Gap	-5.38	-6.42	-11.35	-7.29	-7.07

Note: The total gap in test scores is given by the sample gap $E_W(T(S_T) \mid S_T=s_T)-E_B(T(S_T) \mid S_T=s_T)$.

Table 10
Differences in AFQT Test Score by Level of Schooling at Test Date
Using Structural Model Estimates of Expected AFQT [f=0]

	Control Function Model			Structural Model		
	White	Black	Hispanic	White	Black	Hispanic
Change in Years of Schooling						
Panel A. Males						
μ(ST=10th Grade)-μ(S_T =9th Grade or Less)	13.29	10.45	10.69	17.18	7.80	10.33
μ(ST=11th Grade)- μ(S_T =10th Grade)	4.94	3.34	6.26	6.60	3.85	7.37
μ(ST=High School)- μ(S_T =11th Grade)	3.31	1.10	3.38	3.49	2.72	2.62
μ(ST= 13-14 Years)- μ(S_T =High School)	4.92	3.65	12.32	10.12	6.81	8.39
μ(ST=15 Years or More)- μ(S_T = 13-14 years)	--	--	--	4.74	11.91	--
Panel B. Females						
μ(ST=10th Grade)-μ(S_T =9th Grade or Less)	12.99	11.20	4.74	18.16	9.60	10.10
μ(ST=11th Grade)- μ(S_T =10th Grade)	1.92	-0.77	3.48	2.62	0.60	5.27
μ(ST=High School)- μ(S_T =11th Grade)	2.87	4.21	0.64	2.66	5.21	0.26
μ(ST=13-14 years)- μ(S_T =High School)	7.15	-2.75	0.52	10.54	3.67	4.99
μ(ST=15 years or more)- μ(S_T =13-14 years)	--	--	--	6.19	7.94	--

Notes. The estimates from the structural estimates control for the following covariates: urban status, broken home and south residence at age 14, number of sibling and family income in 1979, mother's and father's education and age at

Table 11

Marginal Effect on AFQT Scores of a Standard Deviation Increase in Latent Ability

Schooling at Test Date	White	Black	Hispanic
Panel A. Males			
S_T=9th Grade or Less	13.56	10.17	10.70
	(0.92)	(0.99)	(1.00)
S_T=10th Grade	17.68	13.98	14.10
	(1.64)	(2.12)	(2.10)
S_T=11th Grade	16.26	15.28	17.68
	(1.60)	(2.12)	(2.62)
S_T=High School	16.58	16.68	16.64
	(1.38)	(2.26)	(2.27)
S_T=13-14 Years	14.55	17.49	17.11
	(1.52)	(2.88)	(3.20)
S_T=15 Years and More	11.91	16.57	--
	(1.84)	(4.67)	--
Panel B. Females			
S_T=9th Grade or Less	13.64	9.27	12.43
	(1.13)	(0.92)	(0.98)
S_T=10th Grade	15.17	12.91	15.23
	(1.87)	(1.60)	(2.24)
S_T=11th Grade	15.52	12.67	15.87
	(1.94)	(1.70)	(2.37)
S_T=High School	14.72	14.53	14.31
	(1.65)	(1.64)	(1.83)
S_T=13-14 Years	13.64	14.42	13.96
	(1.68)	(1.92)	(2.40)
S_T=15 Years and More	12.12	17.15	--
	(1.88)	(3.62)	--

Notes. The estimates from the structural estimates control for the following covariates: urban status, broken home and south residence at age 14, number of sibling and family DQ is a dummy, which is equal to 1 if the school quality of the respondent is higher than tl Where model 1: $T(S_T)=X\beta(S_T)+\lambda(S_T)f+\varepsilon(S_T)$. Model 2: $T(S_T)=X\beta(S_T)+\delta(S_T)Q(S_T)+\lambda(S_T)f+\varepsilon(S_T)$. Model 3:

Table 12

Estimates of AFQT Score Conditional on Factor, E(AFQT | F(factor)=ζ)

Including Faculty per Student as School Quality Covariate

Males From NLSY79

	Whites				Blacks				Hispanics			
	Model 1	Model 2	Model 3		Model 1	Model 2	Model 3		Model 1	Model 2	Model 3	
Schooling at Test Date	Baseline	Linear schoo quality	Low school quality	High school quality	Baseline	Linear schoo quality	Low school quality	High school quality	Baseline	Linear schoo quality	Low school quality	High school quality
						ζ=2.5th percentile						
S_T=9th Grade or Less	19.40	20.62	23.13	18.73	16.79	18.26	19.93	19.76	18.54	21.45	22.94	18.72
	(0.03)	(0.04)	(0.04)	(0.15)	(0.02)	(0.03)	(0.03)	(0.12)	(0.03)	(0.04)	(3.81)	(6.61)
S_T=10th Grade	28.00	24.80	23.20	28.01	18.28	15.07	13.82	17.73	22.47	19.78	24.79	15.03
	(0.03)	(0.04)	(0.05)	(0.09)	(0.02)	(0.04)	(0.04)	(0.08)	(0.04)	(0.06)	(5.46)	(7.40)
S_T=11th Grade	37.56	34.93	35.31	31.88	19.98	19.83	18.02	21.06	23.09	20.49	15.63	28.21
	(0.03)	(0.04)	(0.04)	(0.08)	(0.02)	(0.03)	(0.04)	(0.08)	(0.05)	(0.06)	(6.91)	(6.31)
S_T=High School	40.38	40.09	39.97	38.68	20.38	19.55	18.98	18.65	27.67	26.14	19.79	29.77
	(0.02)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(5.84)	(5.30)
S_T=13-14 years	54.73	52.67	54.48	40.26	25.84	21.95	22.57	22.36	35.16	23.67	26.79	14.20
	(0.03)	(0.04)	(0.05)	(0.09)	(0.05)	(0.08)	(0.09)	(0.13)	(0.08)	(0.14)	(15.03)	(12.14)
S_T=15 Years and More	64.96	64.32	71.08	51.17	39.27	--	--	--	--	--	--	--
	(0.05)	(0.05)	(0.07)	(0.15)	(0.11)	--	--	--	--	--	--	--
						ζ=50th percentile						
S_T=9th Grade or Less	48.70	48.46	45.60	54.45	32.44	32.71	33.19	31.76	38.37	42.14	41.91	45.84
	(0.02)	(0.02)	(0.02)	(0.05)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(2.23)	(2.77)
S_T=10th Grade	66.19	65.71	64.10	69.46	39.78	40.51	40.47	40.76	48.61	50.09	48.54	52.71
	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(2.07)	(2.86)
S_T=11th Grade	72.68	71.60	71.22	72.04	43.48	44.97	44.81	45.87	55.87	57.32	55.67	57.69
	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(2.82)	(2.47)
S_T=High School	76.19	76.84	76.14	79.04	46.03	45.96	46.54	46.00	58.52	60.25	54.78	63.69
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(2.07)	(1.76)
S_T=13-14 Years	86.16	85.84	86.57	82.01	52.74	51.82	54.28	50.07	66.90	63.71	63.73	60.80
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)	(5.83)	(4.88)
S_T=15 Years and More	90.69	90.99	93.37	86.56	64.76	--	--	--	--	--	--	--
	(0.02)	(0.02)	(0.02)	(0.04)	(0.07)	--	--	--	--	--	--	--
						ζ=97.5th percentile						
S_T=9th Grade or Less	71.91	70.28	63.70	83.21	56.50	52.17	51.03	47.88	60.05	62.65	60.87	72.95
	(0.03)	(0.03)	(0.04)	(0.12)	(0.04)	(0.04)	(0.05)	(0.18)	(0.04)	(0.05)	(4.66)	(8.35)
S_T=10th Grade	96.44	97.78	97.02	102.83	72.82	74.76	76.29	71.72	77.18	80.14	72.28	90.38
	(0.02)	(0.03)	(0.04)	(0.07)	(0.05)	(0.05)	(0.06)	(0.14)	(0.05)	(0.06)	(5.63)	(8.35)
S_T=11th Grade	100.51	100.34	100.13	104.38	79.59	78.81	80.82	79.22	91.71	93.82	95.69	87.16
	(0.02)	(0.03)	(0.03)	(0.06)	(0.04)	(0.05)	(0.06)	(0.10)	(0.06)	(0.06)	(7.75)	(6.92)
S_T=High School	104.56	105.64	105.26	111.55	85.45	81.51	83.60	82.76	92.25	94.05	89.77	97.59
	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)	(0.04)	(5.81)	(5.39)
S_T=13-14 years	111.05	111.85	112.41	115.63	94.09	92.04	96.90	87.32	101.60	103.39	100.66	107.39
	(0.02)	(0.02)	(0.03)	(0.05)	(0.05)	(0.07)	(0.08)	(0.12)	(0.06)	(0.08)	(8.42)	(7.94)
S_T=15 Years and More	111.08	111.89	111.32	115.05	103.94	--	--	--	--	--	--	--
	(0.03)	(0.03)	(0.03)	(0.08)	(0.09)	--	--	--	--	--	--	--

Notes. The estimates from the structural estimates control for the following covariates for the X's matrix: urban status, broken home and south residence at age 14, number of siblings, family income in 1979, mother's and father's education, and age at December 1980

Where Q is the percentage of full-time equivalent teachers per student in the high school in which the respondent studied. D_Q is a dummy, which is equal to 1(High Quality of School) if the school quality of the respondent is higher than the average of the sample.

Standard errors are bootstrapped to account for the fact that the mean is an estimated object.

Table 13

Effect of Schooling on AFQT Score Conditional on Factor, E(AFQT | F(factor)=ζ)

Including Faculty per Student as School Quality Covariate

Females From NLSY79

	Whites				Blacks				Hispanics			
	Model 1	Model 2	Model 3		Model 1	Model 2	Model 3		Model 1	Model 2	Model 3	
	Baseline	Linear schoo	Low school	High school	Baseline	Linear schoo	Low school	High school	Baseline	Linear schoo	Low school	High school
Schooling at Test Date		quality	quality	quality		quality	quality	quality		quality	quality	quality
	ζ=2.5th percentile											
S_T=9th Grade or less	22.07	23.20	25.30	17.30	18.23	21.25	21.27	29.88	18.35	19.30	18.58	19.20
	(0.03)	(0.04)	(0.04)	(0.09)	(0.02)	(0.03)	(0.04)	(0.09)	(0.02)	(0.03)	(0.06)	(0.10)
S_T=10th Grade	37.39	38.22	35.72	38.65	21.92	19.72	21.25	15.68	23.83	20.31	16.04	24.99
	(0.03)	(0.03)	(0.05)	(0.05)	(0.02)	(0.03)	(0.04)	(0.06)	(0.03)	(0.05)	(0.07)	(0.08)
S_T=11th Grade	39.37	39.37	38.26	36.44	22.91	21.51	21.36	16.91	28.04	27.57	22.90	27.14
	(0.02)	(0.03)	(0.04)	(0.07)	(0.02)	(0.04)	(0.04)	(0.08)	(0.03)	(0.05)	(0.07)	(0.09)
S_T=High School	43.50	43.62	42.63	42.29	25.10	22.67	22.06	21.83	30.88	27.73	24.89	29.32
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)	(0.06)
S_T=13-14 years	56.04	55.18	53.02	56.44	28.95	27.04	27.40	21.61	36.45	35.91	29.78	76.21
	(0.02)	(0.03)	(0.04)	(0.04)	(0.03)	(0.05)	(0.06)	(0.09)	(0.04)	(0.07)	(0.08)	(0.35)
S_T=15 years and more	65.06	65.80	64.15	63.88	32.45	--	--	--	--	--	--	--
	(0.03)	(0.04)	(0.05)	(0.09)	(0.08)	--	--	--	--	--	--	--
	ζ=50th percentile											
S_T=9th Grade or less	47.09	46.30	41.31	53.72	31.94	32.78	32.68	30.71	37.30	39.37	39.50	39.81
	(0.02)	(0.03)	(0.03)	(0.06)	(0.02)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)
S_T=10th Grade	65.22	66.61	65.29	68.67	41.00	43.50	45.11	41.00	47.04	49.41	48.92	50.50
	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
S_T=11th Grade	67.83	69.26	70.24	67.36	41.64	43.75	43.29	45.13	52.23	53.64	52.53	56.58
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
S_T=High School	70.51	71.07	71.23	70.48	46.58	47.62	47.16	48.72	52.69	54.38	54.78	55.67
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)
S_T=13-14 years	81.07	81.56	80.09	84.30	50.26	51.26	51.48	50.19	57.73	61.00	58.50	74.69
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.08)
S_T=15 years and more	87.29	88.59	88.15	90.18	57.80	--	--	--	--	--	--	--
	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)	--	--	--	--	--	--	--
	ζ=97.5th percentile											
S_T=9th Grade or less	73.31	69.57	57.88	91.40	53.33	47.54	47.04	31.76	66.92	65.08	64.06	64.01
	(0.04)	(0.05)	(0.05)	(0.13)	(0.03)	(0.05)	(0.06)	(0.19)	(0.05)	(0.06)	(0.08)	(0.15)
S_T=10th Grade	94.38	95.21	95.87	99.73	70.78	73.93	75.14	72.85	83.32	86.71	87.51	80.44
	(0.03)	(0.03)	(0.04)	(0.06)	(0.03)	(0.06)	(0.06)	(0.09)	(0.06)	(0.08)	(0.11)	(0.10)
S_T=11th Grade	97.67	99.37	103.32	99.35	70.87	72.22	70.90	80.64	90.04	87.06	87.30	91.14
	(0.03)	(0.03)	(0.05)	(0.06)	(0.03)	(0.05)	(0.07)	(0.10)	(0.05)	(0.06)	(0.10)	(0.11)
S_T=High School	98.81	98.72	100.81	99.64	80.09	79.55	78.73	82.55	86.79	88.53	89.87	86.61
	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.06)	(0.04)	(0.05)	(0.06)	(0.07)
S_T=13-14 years	107.29	108.14	108.10	113.11	83.52	82.27	81.77	86.15	90.99	93.15	92.21	72.92
	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)	(0.07)	(0.05)	(0.06)	(0.08)	(0.27)
S_T=15 years and more	110.58	111.55	112.98	117.37	97.36	--	--	--	--	--	--	--
	(0.03)	(0.03)	(0.04)	(0.07)	(0.08)	--	--	--	--	--	--	--

Notes. The estimates from the structural estimates control for the following covariates for the X's matrix: urban status, broken home and south residence at age 14, number of siblings, family income in 1979, mother's and father's education, and age at December 1980

Where Q is the percentage of full-time equivalent teachers per student in the high school in which the respondent studied. D_Q is a dummy, which is equal to 1(High Quality of School) if the school quality of the respondent is higher than the average of the sample.

Standard errors are bootstrapped to account for the fact that the mean is an estimated object.

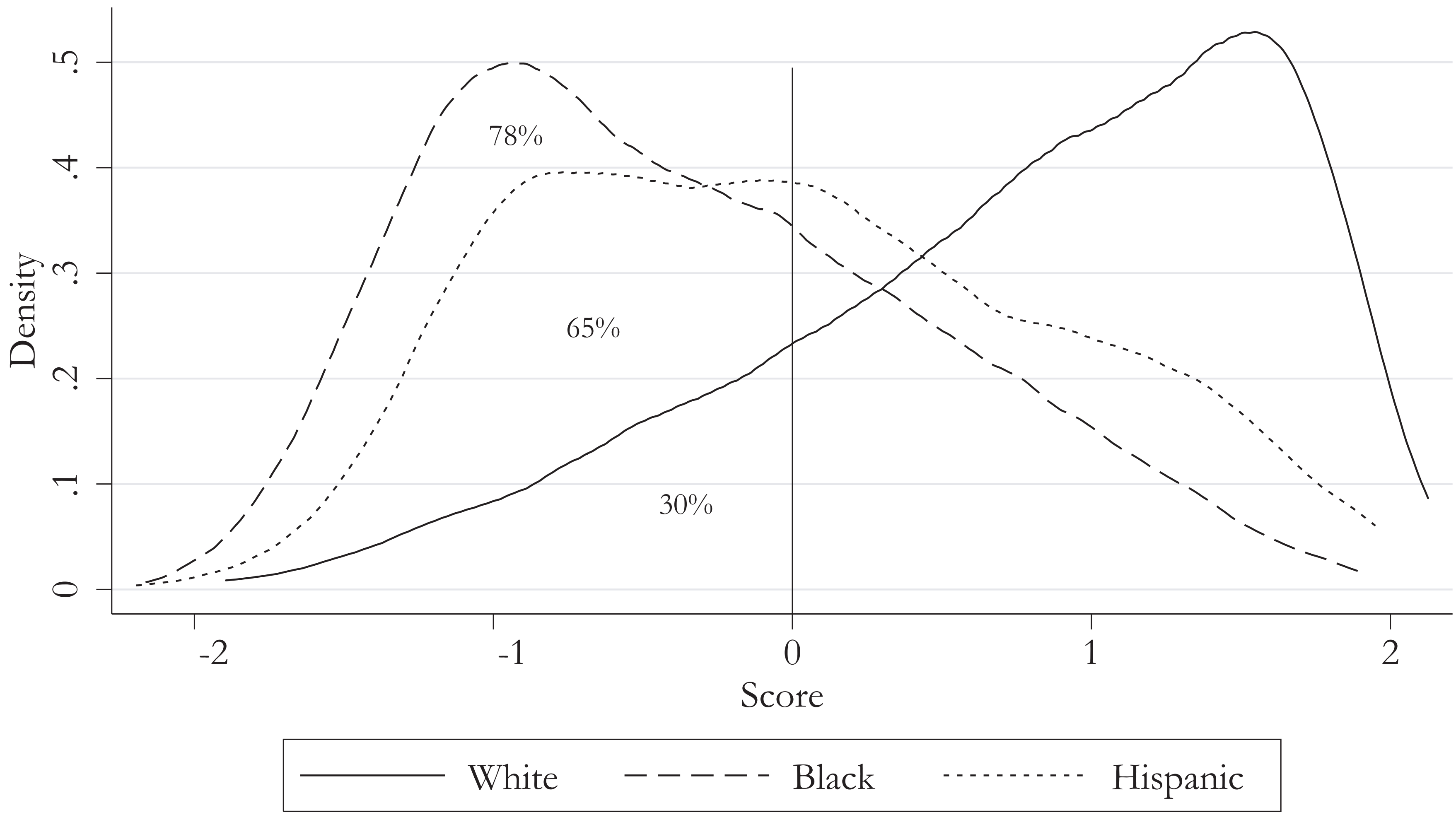
Table A1

Sample Size by Final School Level and Schooling at Test Date, NLSY79

	Highest Grade Completed at Test Date, S_T						Total
Final Schooling, S	Less than or Equal to 9	10th Grade	11th Grade	High School	13-14 Years	15 Years or More	
Panel A. White Males							
Dropout & GED	133	93	67	52	--	--	345
High School	32	104	98	483	--	--	717
Some College & 2-Year Degrees	15	57	51	153	127	10	413
4 Years of College or More	14	79	88	109	162	136	588
Total	194	333	304	797	289	146	2063
Panel B. White Females							
Dropout & GED	99	74	51	55	--	--	279
High School	23	96	100	532	--	--	751
Some College & 2-Year Degrees	12	70	62	173	177	12	506
4 years of College or More	4	72	68	115	198	144	601
Total	138	312	281	875	375	156	2137
Panel C. Black Males							
Dropout & GED	116	105	78	42	--	--	341
High School	45	72	74	264	--	--	455
Some College & 2-Year Degrees	15	34	37	79	81	3	249
4 Years of College or More	3	15	31	26	40	34	149
Total	179	226	220	411	121	37	1194
Panel D. Black Females							
Dropout & GED	85	71	44	32	--	--	232
High School	23	64	67	252	--	--	406
Some College & 2-Year Degrees	15	45	53	134	148	11	406
4 Years of College or More	0	23	22	36	73	37	191
Total	123	203	186	454	221	48	1235
Panel E. Hispanic Males							
Dropout & GED	103	65	50	20	--	--	238
High School	25	34	40	128	--	--	227
Some College & 2-Year Degrees	10	25	26	54	56	2	173
4 Years of College or More	4	13	14	11	21	18	81
Total	142	137	130	213	77	20	719
Panel F. Hispanic Females							
Dropout & GED	101	50	28	17	--	--	196
High School	14	38	39	143	--	--	234
Some College & 2-Year Degrees	17	35	42	73	68	7	242
4 Years of College or More	2	14	16	24	28	19	103
Total	134	137	125	257	96	26	775

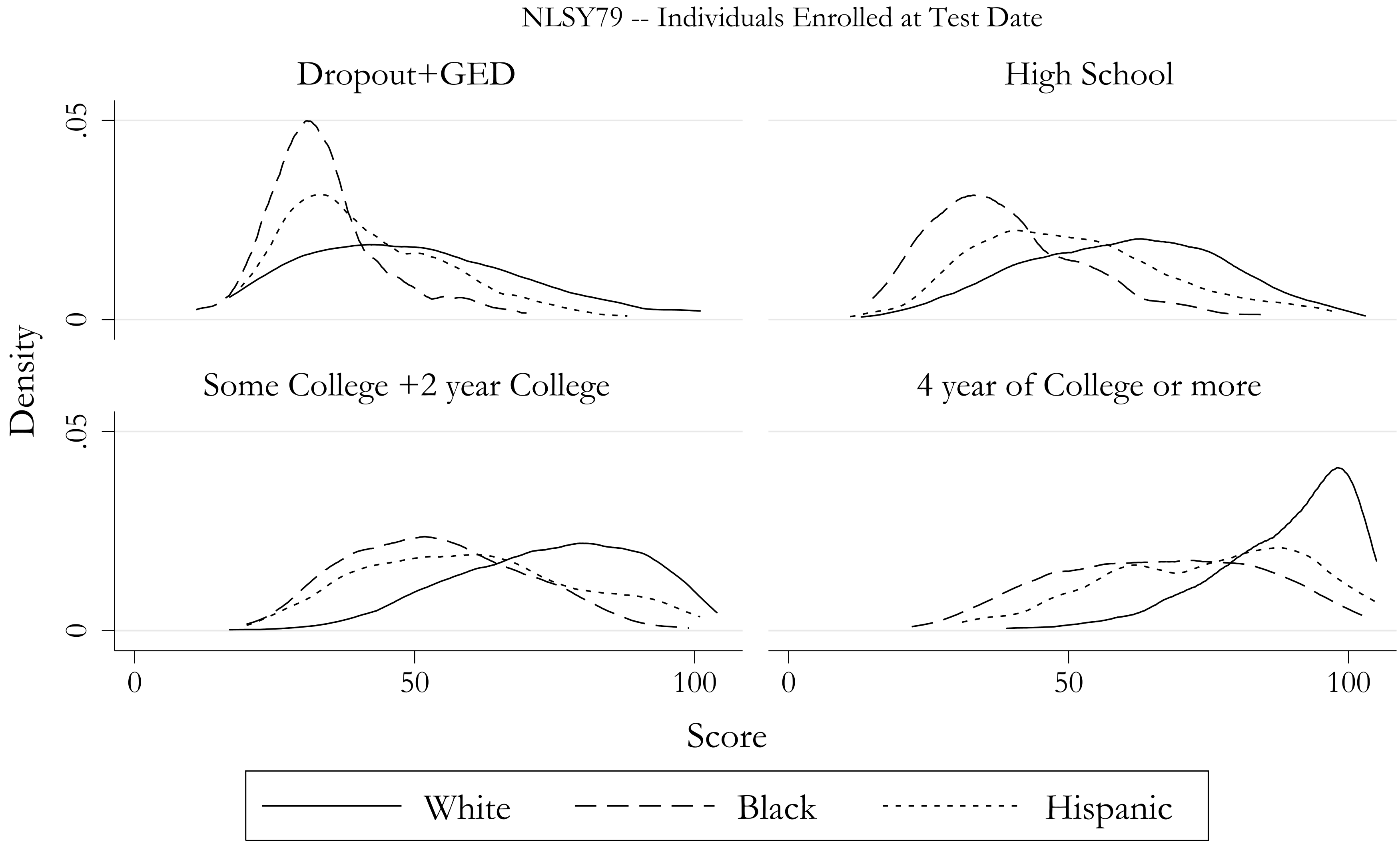
Note: The character "--" defines unavailable cells. GED recipients are defined as those with 12 years of completed schooling at testing date

Figure 1
Density of Age-Corrected AFQT
NLSY79 -- Individuals Enrolled at Test Date



Note. AFQT is a subset of 4 out of 10 ASVAB tests used by the military for enlistment screening and job assignment. It is the aggregate score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests. Age adjusted AFQT is the standardized residual from the regression of the AFQT score on age at the time of the test dummy variables.

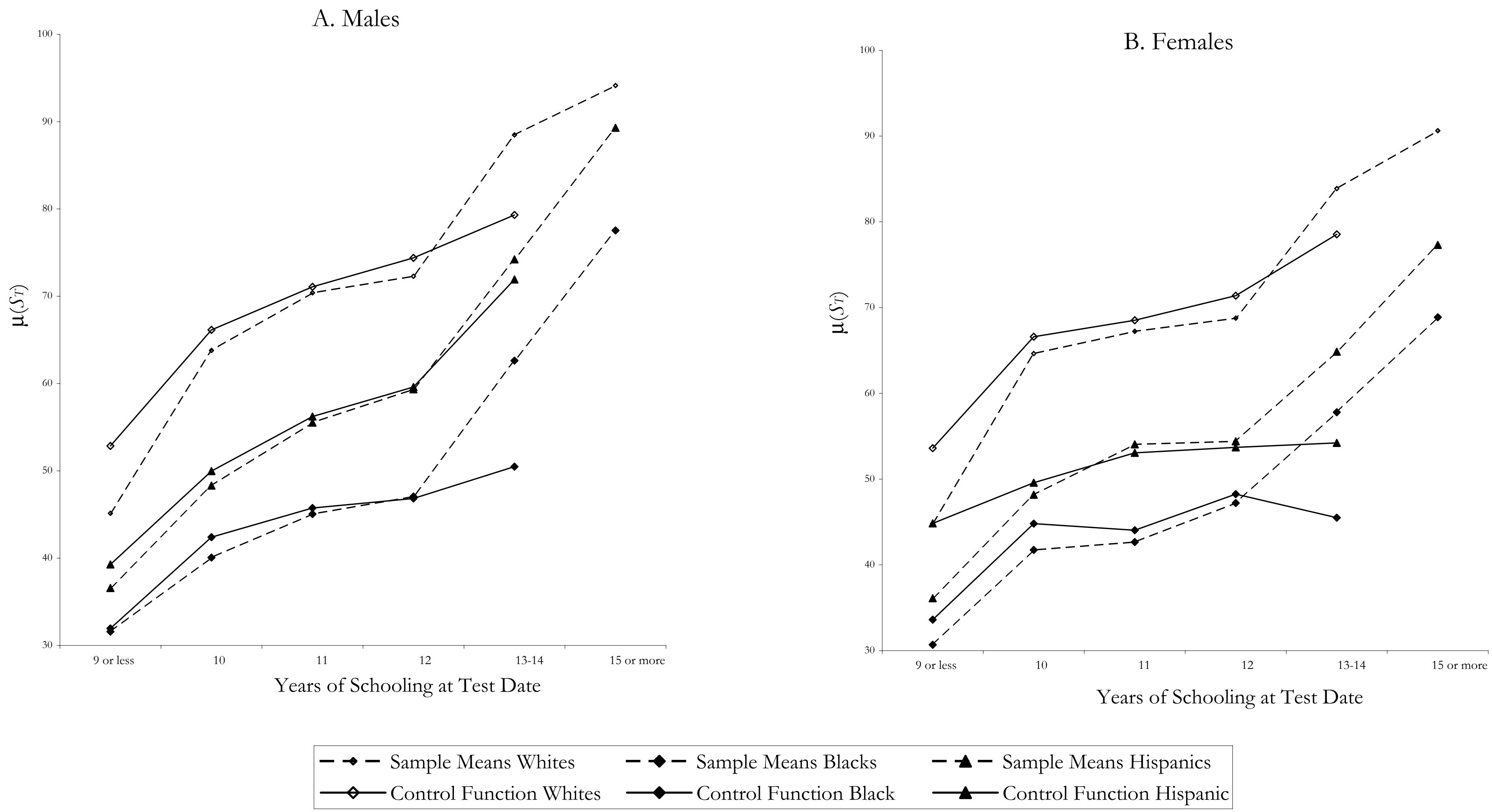
Figure 2
Density of AFQT by Final Schooling Level



Note. Education corresponds to final schooling. AFQT is a subset of 4 out of 10 ASVAB tests used by the military for enlistment screening and job assignment. It is the aggregated score from the word knowledge, paragraph comprehension, mathematics knowledge, and arithmetic reasoning ASVAB tests.

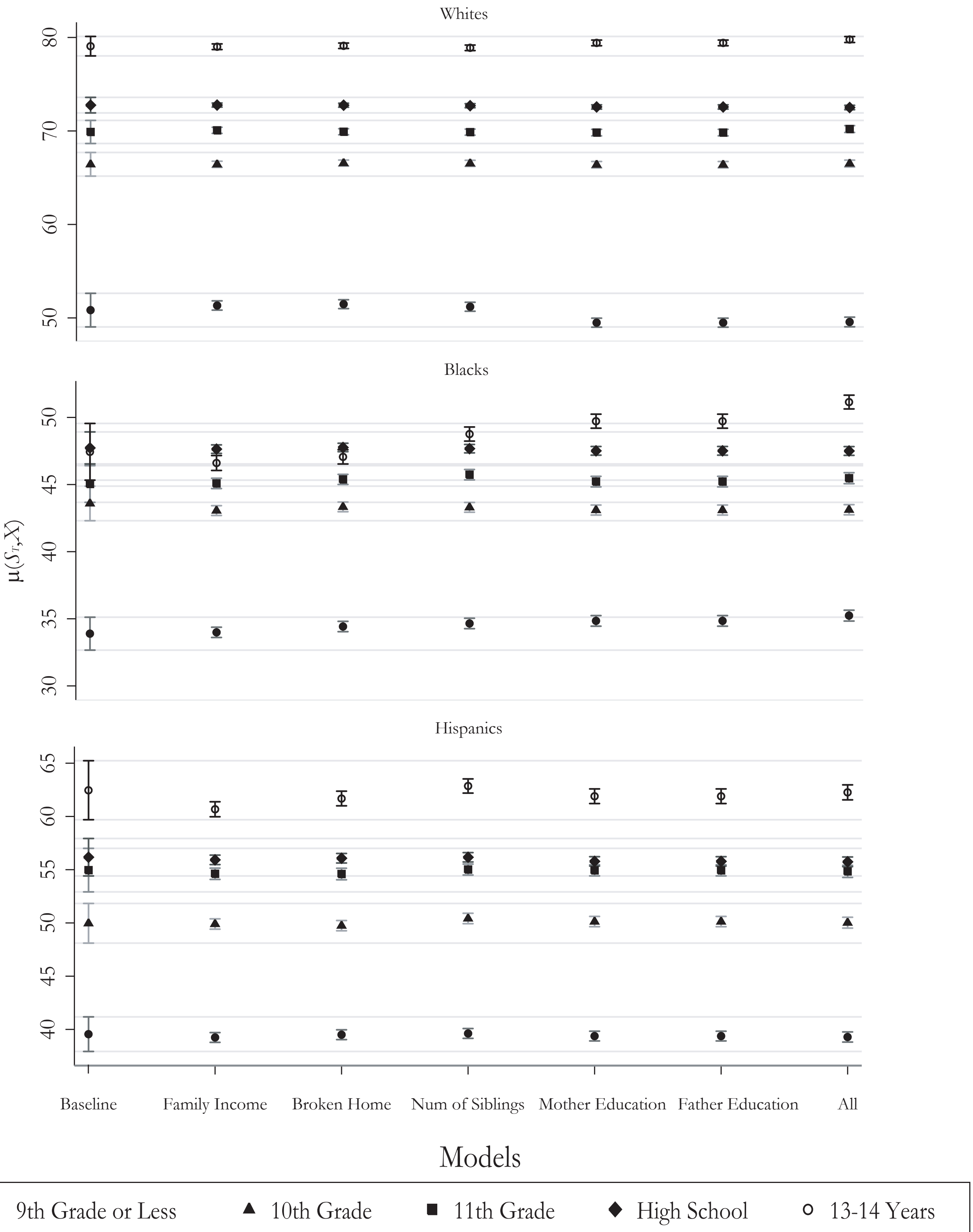
Figure 3

Comparison of Sample Means of Test Scores and Control Function Estimates [*f*=0]



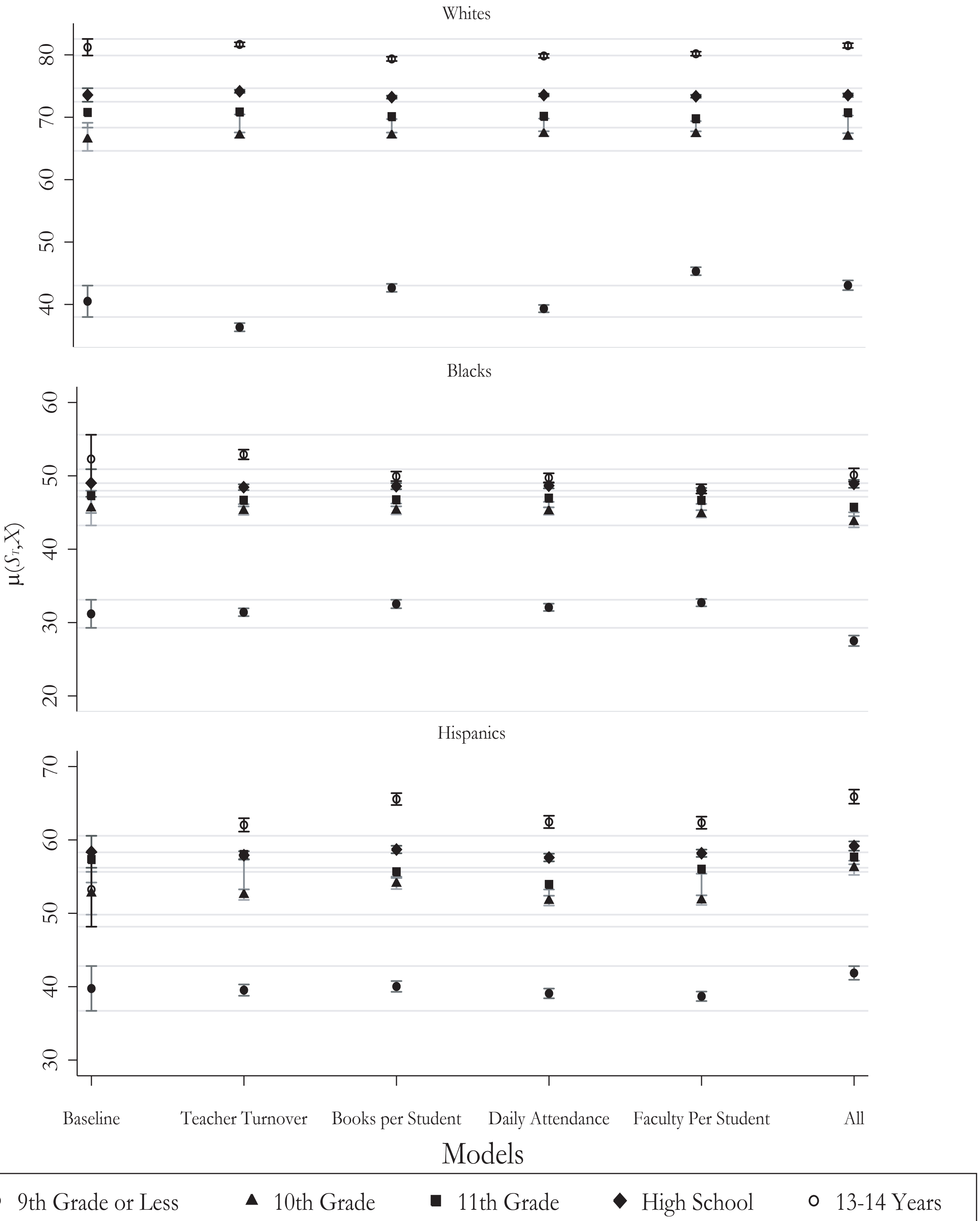
Notes. The estimates, μ(ST), from the control function approach correspond to the expected test score when *f*=0.

Figure 4
Control Function Estimates of Expected Test Score [f=0] Including Family Background Variables by Schooling at Test Date



Notes. The estimates of represent the estimates $\mu(ST,X)$, from the control function model. The baseline model does not include covariates in X. The different models include in the vector X a gender dummy and different family background variables. The estimates include the following variables, one at a time: family income in 1979, broken home at 14, number of siblings at 14, mother and father education. Finally all these family background variables are included. The whiskers represent the 95 percent confidence interval of each estimate.

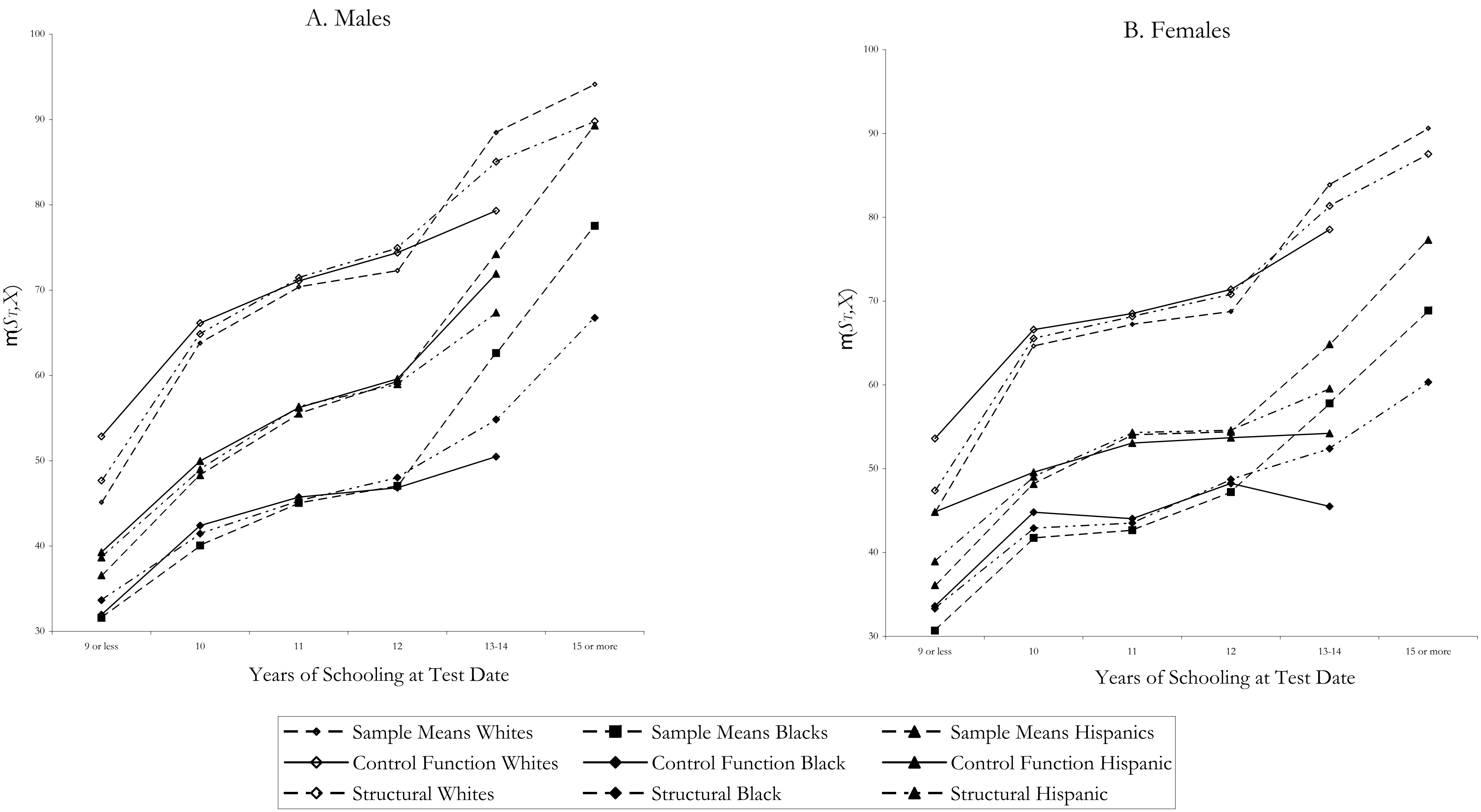
Figure 5
Control Function Estimates of Expected Test Score [f=0] Including School Quality Variables
by Schooling at Test Date



Notes. The estimates represent the estimates μ(ST,X), from the control function model. The baseline model does not include covariates in X. The different models include in the vector X a gender dummy and different school quality variables. The estimates include the following variables, one at a time: teacher turnover excluding death or retirement, books per student, daily percentage attendance in school and faculty per student. Finally all these school quality variables are included, the sample includes only individuals with data on the four school quality characteristics. The baseline model is restricted to this last sample.The whiskers represent the 95 percent confidence interval of each estimate.

Figure 6

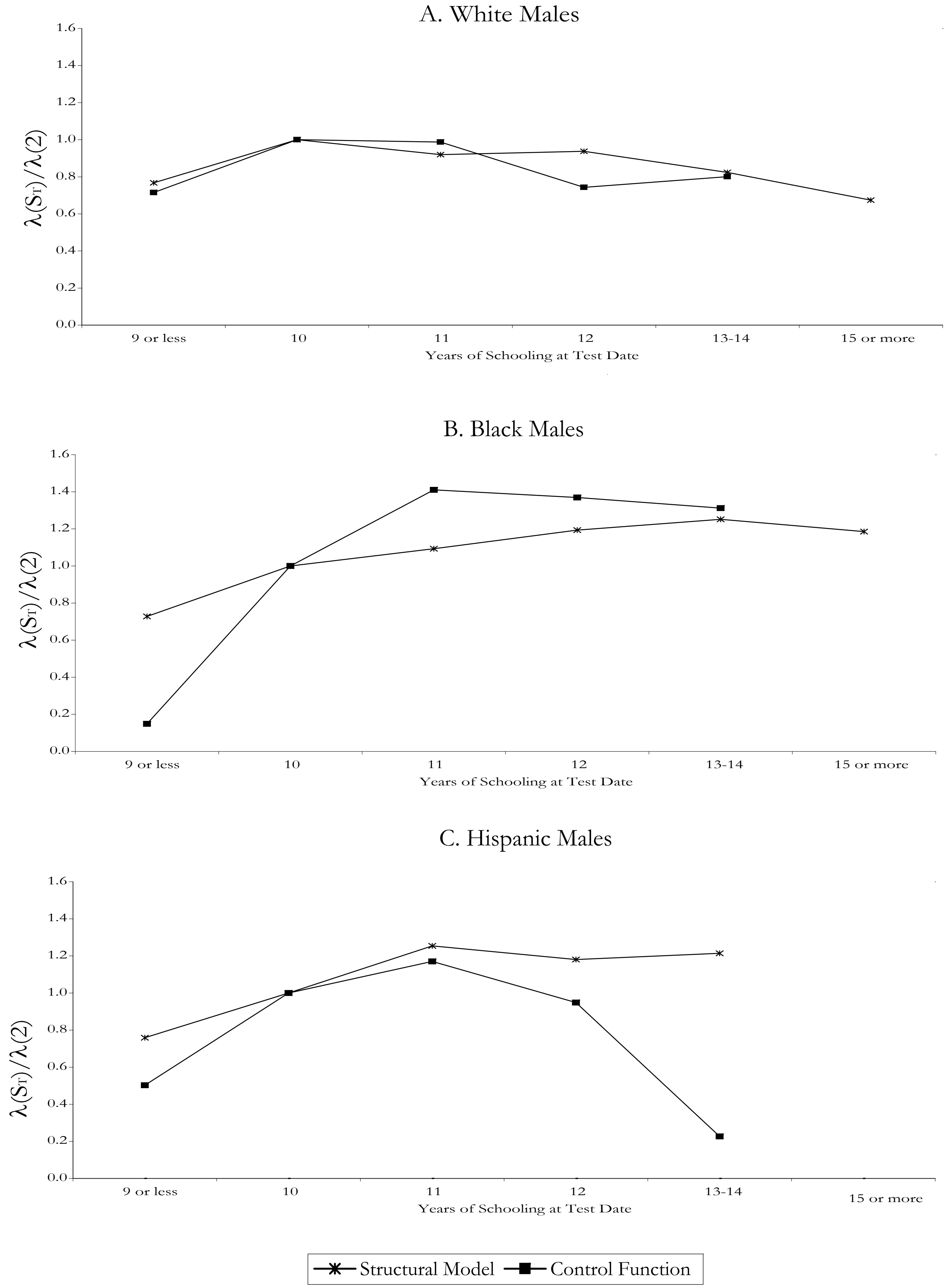
Comparison of Sample Means, Control Function Estimates and Structural Estimates of Expected Test Score [f=0]



Notes: The estimates from the control function presented do not include covariates X. The dashed lines correpond to the estimated test scores from the structural model controlling for covariates X when f=0. The estimates from the structural estimates control for the following covariates: urban status, broken home and southern residence at age 14, number of siblings and family income in 1979, mother's and father's education, and age at December 1980.

Figure 7

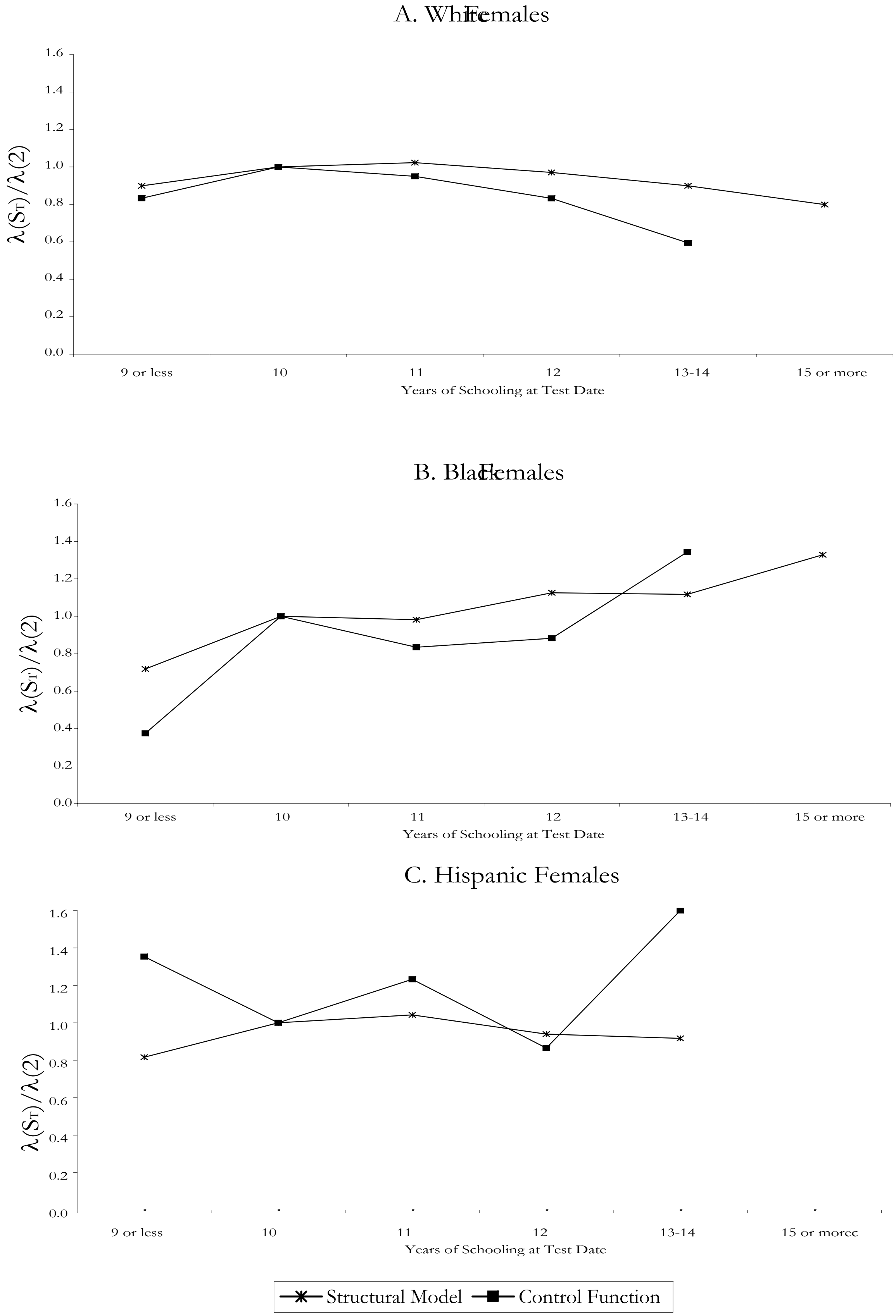
Comparison of Control Function and Structural Estimates of Ratios of Factor Loadings



Notes: The estimates from the control function presented do not include covariates X. The estimated from the structural model control for the following covariates: urban status, broken home and south residence at age 14, number of siblings and familiy income in 1979, mother's and father's education, and age at December 1980.

Figure 8

Comparison of Control Function and Structural Estimates of Ratios of Factor Loadings

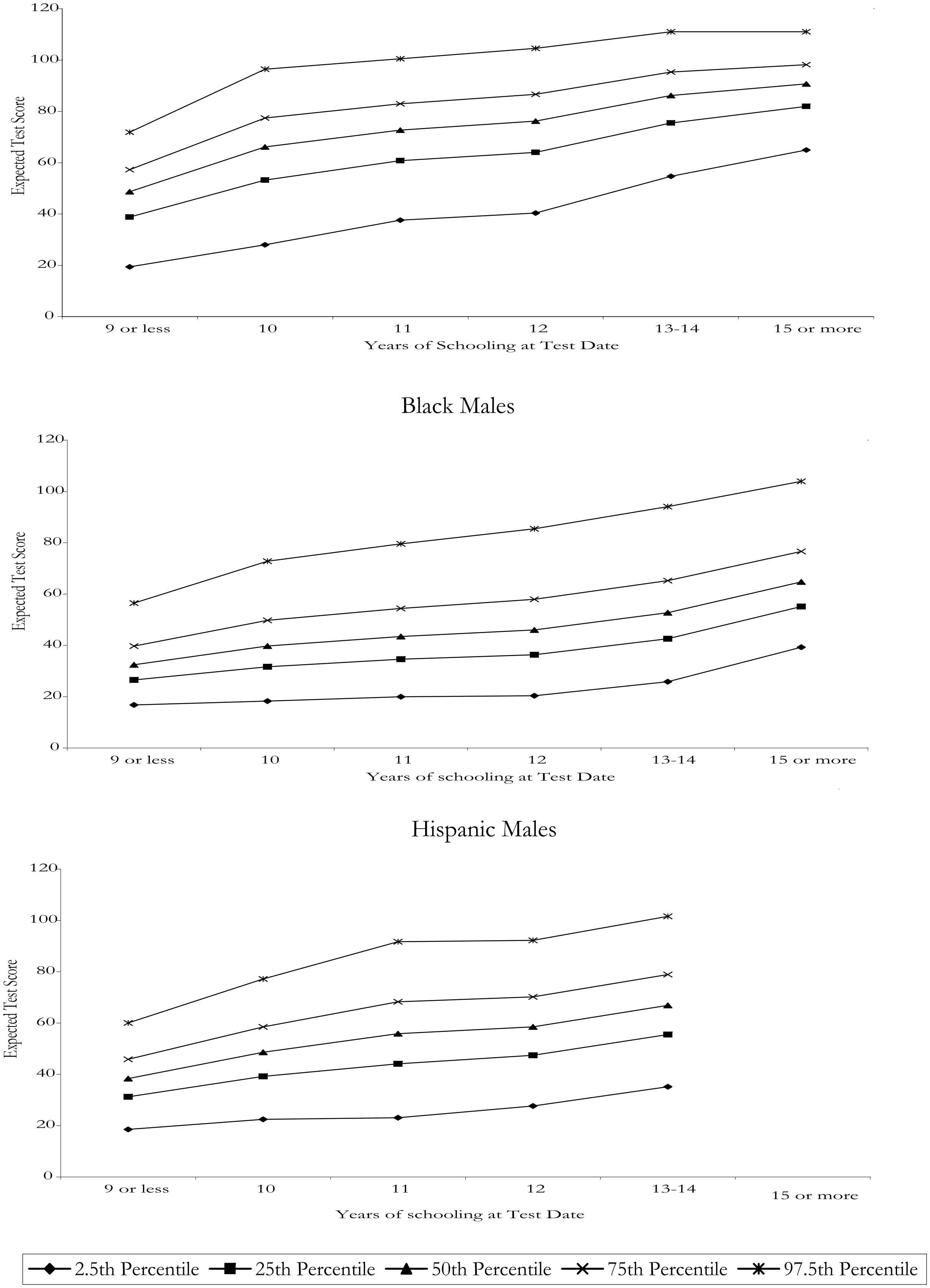


Notes: The estimates from the control function presented do not include covariates X. The estimated from the structural model control for the following covariates: urban status, broken home and south residence at age 14, number of siblings and familiy income in 1979, mother's and father's education, and age at December 1980.

Figure 9

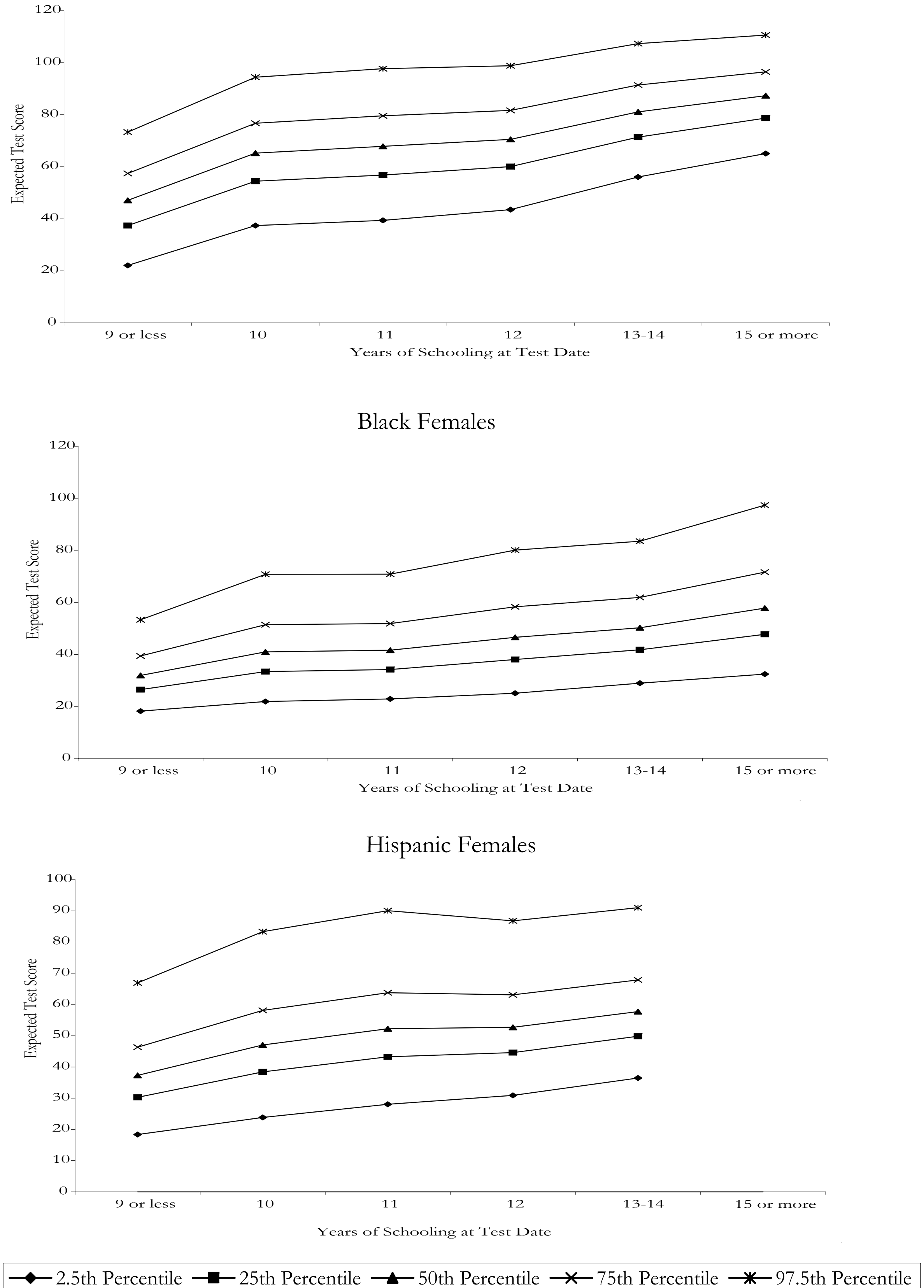
Expected Value of AFQT Score Conditional on Factor

White Males



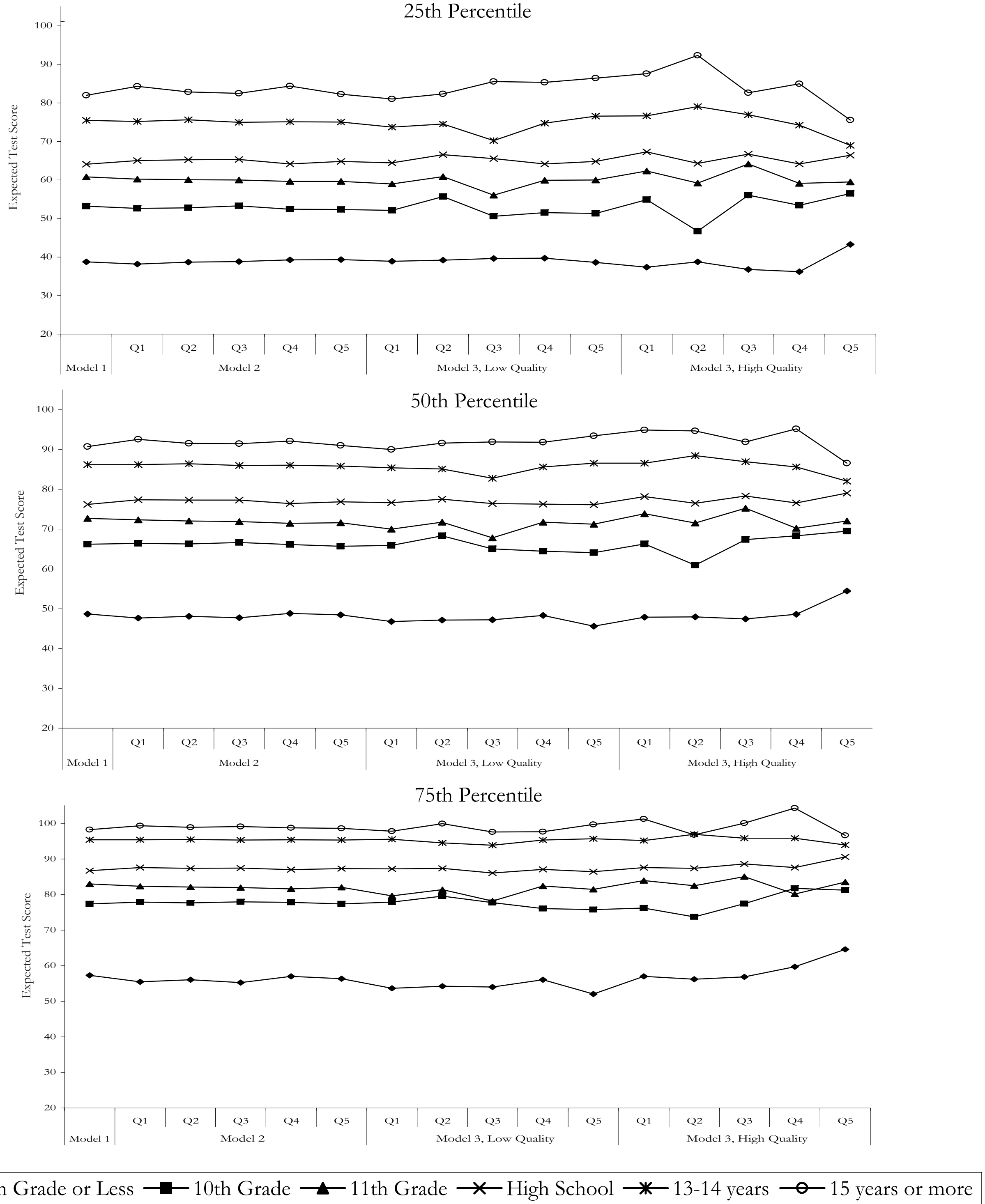
Notes: The estimated from the structural model include the follwong covariates: urban status, broken home and southern residence at age 14, number of siblings and familiy income in 1979, mother's and father's education, and age at December 1980.

Figure 10
Expected Value of AFQT Score Conditional on Factor
White Females



Notes. The estimates from the structural estimates inlude the following covariates: urban status, broken home and south residence at age 14, number of sibling, family income in 1979, mother's and father's education, and age at December 1980.

Figure 11
Expected AFQT Score Conditional on Factor Level Including School Quality
White Males from NLSY 79

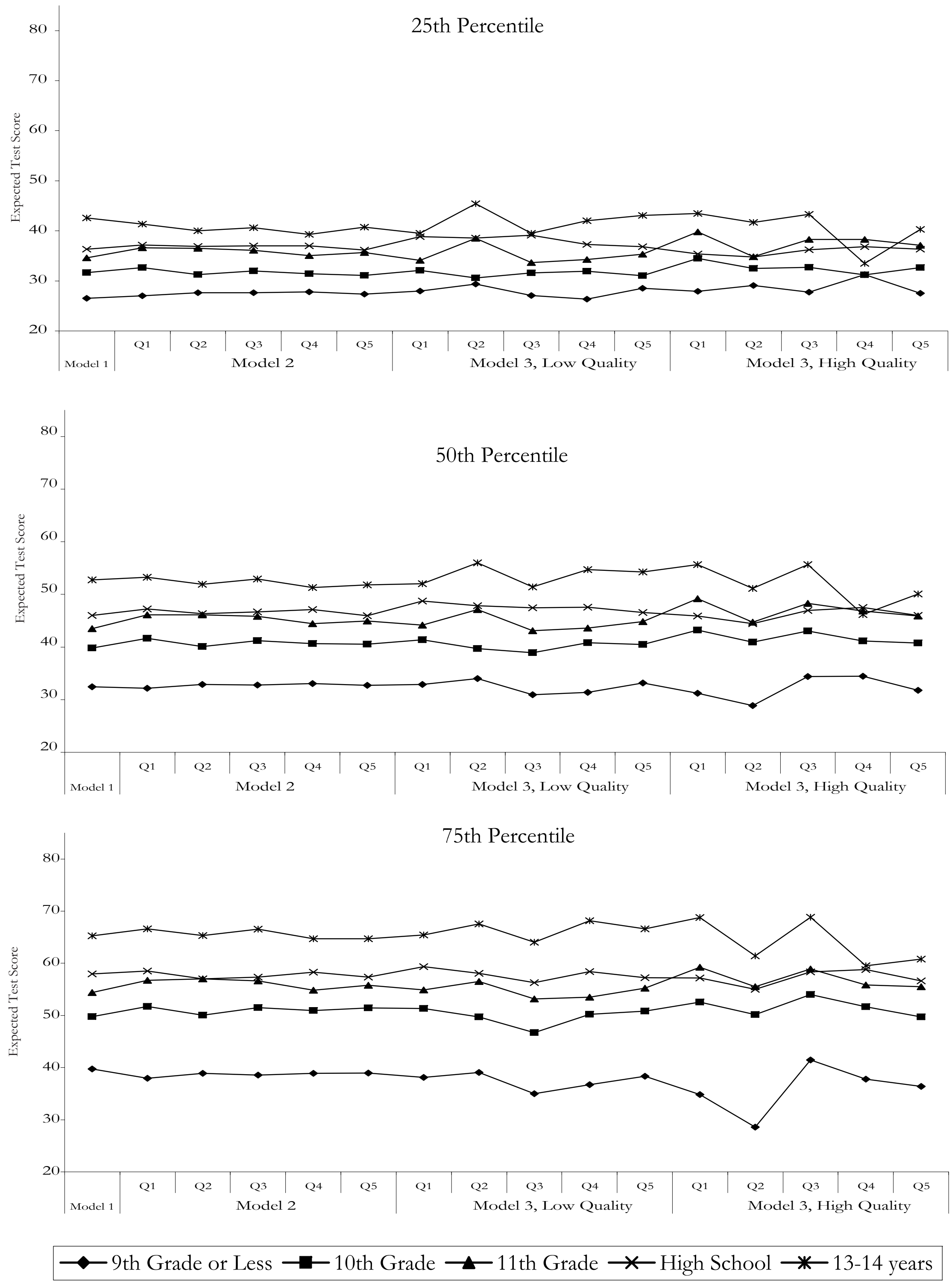


Notes: Model 1 does not include school quality variables, the estimated equation: $T(S_T)=X(S_T)\beta(S_T)+\lambda(S_T)f$. Model 2 includes school quality linearly in the test score equation: $T(S_T)=X(S_T)\beta(S_T)+\delta(S_T)Q(S_T)+\lambda(S_T)f$. Model 3 is estimated for two samples, low and high quality. In particular we estimate: $T(S_{T,}D_Q)=X(S_{T,}D_Q)\beta(S_{T,}D_Q)+\lambda(S_{T,}D_Q)f$, where Q corresponds to a measure of school quality. D_Q is a dummy variable that identifies individuals with high quality, those with quality above the mean, and those with low quality for each of the six demografic groups. The measures of school quality are: Q1=annual salary for a new certified teacher with a BA; Q2=teacher turnover excluding death and retirement; Q3=Percentage of faculty with MA or PhD; Q4=books per student and Q5=faculty per student. In addition, the estimates include the following controls: urban status, broken home, south of residence at age 14, number of sibling and family income in 1979, mother's and father's education, and age at December 31, 1980. The normalization points for the factor loading corresponds in Model1, 2 and 3 to $\lambda(S_T$=9th grade or less)=1 for the word knowledge test. In addition, in Model 3 the estimates are normalized to low school quality.

Figure 12

Expected AFQT Score Conditional on Factor Level Including School Quality

Black Males from NLSY 79

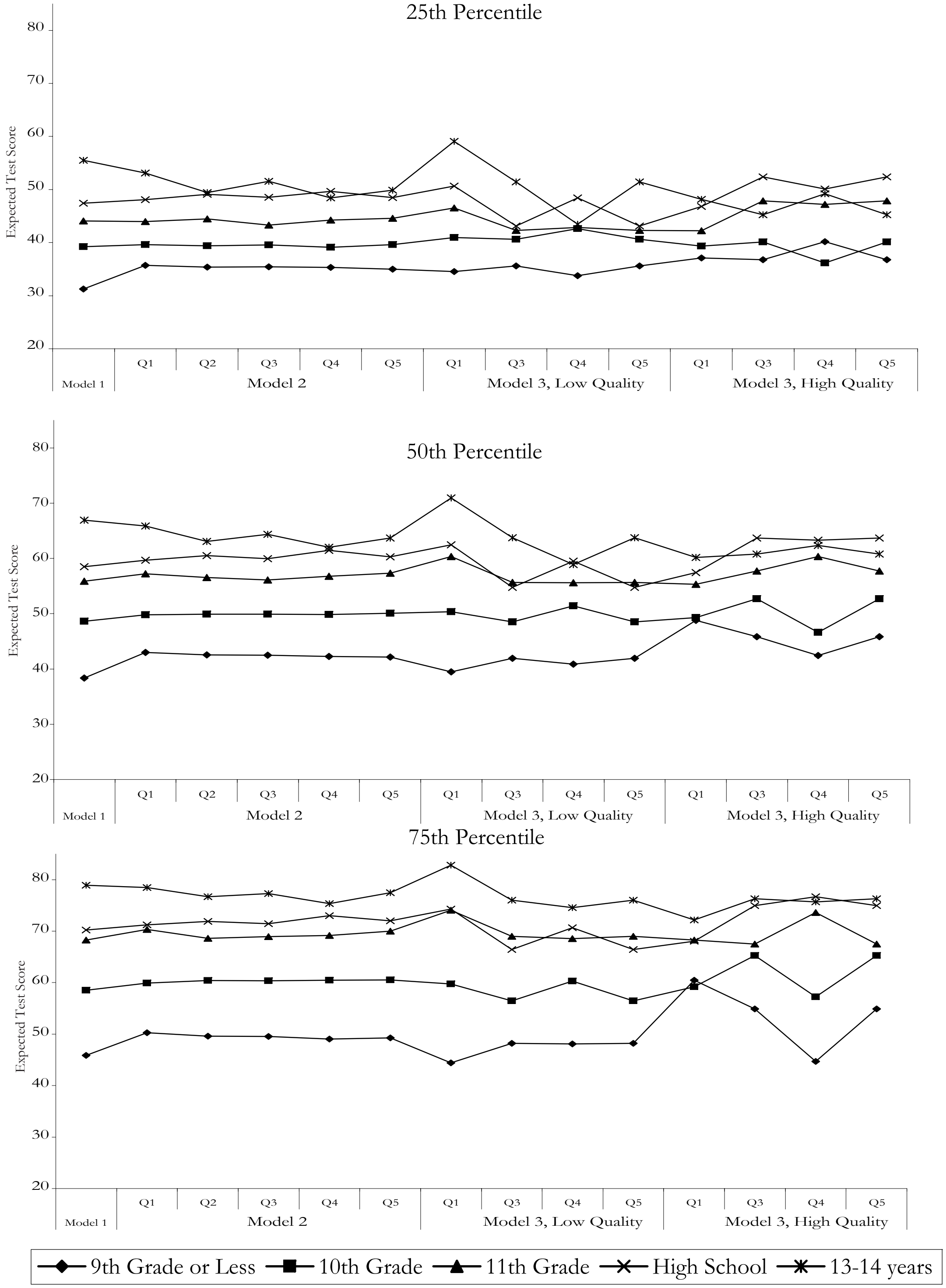


Notes: Model 1 does not include school quality variables, the estimated equation: $T(S_T)=X(S_T)\beta(S_T)+\lambda(S_T)f$. Model 2 includes school quality linearly in the test score equation: $T(S_T)=X(S_T)\beta(S_T)+\delta(S_T)Q(S_T)+\lambda(S_T)f$. Model 3 is estimated for two samples, low and high quality. In particular we estimate: $T(S_T,D_Q)=X(S_T,D_Q)\beta(S_T,D_Q)+\lambda(S_T,D_Q)f$, where Q corresponds to a measure of school quality. D_Q is a dummy variable that identifies individuals with high quality, those with quality above the mean, and those with low quality for each of the six demografic groups. The measures of school quality are: Q1=annual salary for a new certified teacher with a BA; Q2=teacher turnover excluding death and retirement; Q3=Percentage of faculty with MA or PhD; Q4=books per student and Q5=faculty per student. In addition, the estimates include the following controls: urban status, broken home, south of residence at age 14, number of sibling and family income in 1979, mother's and father's education, and age at December 31, 1980. The normalization points for the factor loading corresponds in Model1, 2 and 3 to $\lambda(S_T$=9th grade or less$)=1$ for the word knowledge test. In addition, in Model 3 the estimates are normalized to low school quality.

Figure 13

Expected AFQT Score Conditional on Factor Level Including School Quality

Hispanic Males from NLSY 79



Notes: Model 1 does not include school quality variables, the estimated equation: $T(S_T)=X(S_T)\beta(S_T)+\lambda(S_T)f$. Model 2 includes school quality linearly in the test score equation: $T(S_T)=X(S_T)\beta(S_T)+\delta(S_T)Q(S_T)+\lambda(S_T)f$. Model 3 is estimated for two samples, low and high quality. In particular we estimate: $T(S_T, D_Q)=X(S_T, D_Q)\beta(S_T, D_Q)+\lambda(S_T, D_Q)f$, where Q corresponds to a measure of school quality. D_Q is a dummy variable that identifies individuals with high quality, those with quality above the mean, and those with low quality for each of the six demografic groups. The measures of school quality are: Q1=annual salary for a new certified teacher with a BA; Q2=teacher turnover excluding death and retirement; Q3=Percentage of faculty with MA or PhD; Q4=books per student and Q5=faculty per student. In addition, the estimates include the following controls: urban status, broken home, south of residence at age 14, number of sibling and family income in 1979, mother's and father's education, and age at December 31, 1980. The normalization points for the factor loading corresponds in Model1, 2 and 3 to $\lambda(S_T$=9th grade or less)=1 for the word knowledge test. In addition, in Model 3 the estimates are normalized to low school quality.