# Measurement Errors in Models of the Racial and Ethnic Test Scores Gaps

(Preliminary)

Hongyun Han

Department of Sociology

The University of Wisconsin-Madison

September, 2008

**ABSTRACT**

The measurement errors in children's test scores have not been well understood in the studies of the racial test score gap. This paper aims to address the cumulative racial gap in test scores by taking measurement errors into account. Using a newly available dataset (the Early Childhood Longitudinal Study), we use structural equation models to estimate the effect of family background on the racial gap in test scores among children from kindergarten to fifth grade. We find that the model with measurement error correction and causal chain fits the data best, suggesting that family background and test scores are latent variables, and minority children face cumulative disadvantages during elementary school. Furthermore, our findings show that latent family background has a stronger influence on achievement among African American fifth graders than their White counterparts, after correcting measurement errors in test scores and family background.

## Introduction

The racial gap in achievement test scores has attracted extensive attention from scholars, policy makers and the general public. African American and Hispanic students have persistently achieved lower test scores than their White counterparts (Jencks & Philips, 1998; Ream, 2005), and the racial gap in test scores has widened since 1990 (Philips, 2000; Grissmer et al,2003; Kao & Thompson, 2003). A large number of studies have shown that the substantial lag in test scores in elementary school has a life-long negative influence, including retention between grades, dropping out of high school, unemployment, and lower earnings (Jencks & Philips, 1998). In the journey to closing the racial gap, scholars have proposed that parental socioeconomic status is an important factor in explaining lower achievement among minority students, which is often measured by parental education, occupation and income. Recently, Yeung and Conley (2008) have added another dimension, family wealth, to the list of family background variables.

There are three major problems with the studies examining the association between family background and the racial gap in test scores. First, random and nonrandom measurement errors in respondents' subjective report of family background could bias the measurements. Bielby, Hauser and Featherman (1977) have shown that Blacks' reporting on social background and achievement is subject to errors, and analysis ignoring such an error pattern "exaggerates racial differences in returns to schooling and occupational inequality not attributable to social origins." Hauser and Andrew (2007) contend that "neither students nor their parents are very reliable or precise in reporting key socioeconomic characteristics of the parents" and Hispanic and Black parents have a lower quality of objective reporting than Whites. The substantial racial differentials in the quality of socioeconomic reporting are of paramount importance in the analysis of the association between socioeconomic status and the racial test scores gap. Second, nonrandom measurement errors in test scores could also lead to biased estimates of racial differentials in test scores. Jencks (1998) argues that content bias in a test may lead to a biased estimate, since the test contains questions that favor one group over another. A few studies have shown that the black-white gap was larger for some words than for others in the Peabody Picture Vocabulary test (PPVT) due to pronunciation differences (Jencks 1998). Therefore, there is reason to suspect that studies may report biased estimates if they ignore systematic errors or bias in test scores against particular minority groups. Third, it is almost impossible to exhaust the factors under the title of "family background."[1] So, no matter how long the list of family background could possibly be, critics point out that there are still unobserved factors for which a particular model fails to control. Finally, the growth model is sensitive to measurement errors in modeling the cumulative disadvantages experienced by some of the minority students. Therefore, our understanding about the roots of the racial test scores gap is hindered by often contradictory conclusions on the effect of family background[2].

[1] Sibling models are the exceptions.
[2] For example, two studies using the same ECLS-K data yield very conflicting results in the effect of family background on the black-white test scores in the first two years of elementary school. Fryer and Levitt (2004) conclude that none of family background matter, but Downey, Hippel and Broh (2004) find that family is the source of inequality in summer learning, while children progress equally in school.

In this study, we argue that the measurement errors in children's test scores should be seriously addressed in modeling the racial test scores gap. Using a newly available dataset (the Early Childhood Longitudinal Study), we use structural equation models to estimate the effect of family background on the racial gap in test scores among children from kindergarten to fifth grade. We find that the model with measurement error correction and a causal chain fits the data best, suggesting that family background and test scores are latent variables, and cumulative disadvantages among minority children. Furthermore, our findings show that latent family background has a weaker influence in African American fifth graders.

## Sources of measurement errors in the ECLS-K

The measures of the test scores and parental education and occupation are not free from errors in the ECLS-K. Errors can be introduced in the child's assessment, the parental interview, and the data coding process.

First, measurement errors in parental education stem in part from the data collection procedures. The information on parents' highest education attained was collected in the first two waves (fall and spring semester of kindergarten)[3] in 1998 and 1999. In the later four waves, the ECLS-K only asked the respondents to report any changes in their education level, instead of soliciting full information on the current education level. Such a procedure may lead to serious problems in assessing parental education: a) proxy report

[3] For households not interviewed in fall-kindergarten (e.g., parents of children in refusal-converted schools), this information was collected in spring-kindergarten.

error, either from mothers' report on their spouse's education or from non-parents' report on parents' education; and b) wrong or incomplete reporting of changes in educational level.

Second, the measurement of parental occupation is problematic in the process of data collection and coding. The ECLS-K asks the respondents to report in the fall semester of kindergarten on the parents' occupation in regards to the industry, type of work and job duties, and the parental occupation is not updated as children progress through the grades. During the coding process, the occupation was coded using the "Manual for Coding Industries and Occupations," March 1999 (National Household Education Survey, NHES: 99), with 13 aggregate level industry codes and 22 occupation codes. Almost 75% of the observations were coded by two coders, and a coding supervisor arbitrated any disagreement. The final step was to generate the prestige scores for the occupation by averaging the 1989 General Social Survey (GSS) prestige score of the occupation[4]. Therefore, two major problems arise. First, the coding process may be subject to errors: a) the final decision in a disagreement regarding occupation coding was left to the discretion of one individual, whose credibility in occupation measuring is questionable; b) some of the coders' errors in occupation were unchecked; c) coarser aggregate level prestige scores lose the subtle difference between occupations. Second, the prestige scores may not well capture the effect of occupational standing (Hauser and Warren 1997).

[4] ECLS-K base year user guide, chapter 7

Third, at least two types of practices in direct child assessment may introduce measurement errors in test scores. The first type of practice is to use an Oral Language Development Scale (OLDS) to screen children whose primary home language is not English[5]. The OLDS measures children's listening comprehension, vocabulary, and ability to understand and produce language. Only children who pass the OLDS receive the full version of English test, which make the Hispanic children a selective sample in the ECLS-K. The administrator has full discretion to determine whether a child passes a cut off point in the OLDS scale score, so it is likely to contain the "content bias" argued by Jencks (1998). The second type of practice relates to the completeness of the child's assessment. The assessments were shortened or discontinued "if the administrator perceived that the child was uncomfortable or distressed about responding to the assessment items."[6] So, some of the children scored lower because they failed to complete the tests, again at the discretion of the administrator. Therefore, the measurement errors in the test scores should be taken into account to make meaningful comparisons across racial groups.

## Data

We use the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) to perform the analysis. The ECLS-K is conducted by the Department of Education at children's entry into kindergarten and in their progression through school to provide information on a number of educational outcomes. The nationally representative sample consists of a

[5] ECLS-K base year user guide, chapter 2
[6] ECLS-K base year user guide, chapter 2

cohort of children who entered kindergarten in the fall of 1998. The ECLS-K used a multistage probability design [PSUs-Schools-Students]. The original sample at the fall semester of kindergarten included 21,260 kindergartners from 1,280 schools. Various types of information were collected on the children at six different times in the study (ie, during the fall [1998; wave 1] and spring [1999; wave 2] of the children's kindergarten year, in the fall [1999; wave 3] and spring [2000; wave 4] of their year in first grade, spring [2002; wave 5] of their year in third grade and spring [2004, wave 6] of their year in fifth grade). We use the reading and math IRT test scores of five waves (excluding a 30% subsample in the fall of the first grade), since they are recalibrated in the fifth grade to make consistent comparison between grades.

To produce a correlation matrix of test scores and parental characteristics, our preliminary sample is restricted to children who have test scores across five waves. This restriction yields a sample of 6,846 children, with 5,221 Non-Hispanic Whites, 584 African Americans, and 1041 Hispanics. The African American children in this sample are underrepresented, mainly due to the lack of test scores across five waves, and the larger probability of attrition.

Figures 1-5 in the Appendix show the distribution of the gender-specific IRT reading test scores across five waves in a box plot. The test scores in the fall and spring semesters of kindergarten are highly skewed at the top of the distribution, with larger than 2.0 values of skewness and kurtosis.  As children progress through the grades, the skewness is reduced in the spring of the first and the third grade, yet begins to skew slightly toward the bottom of the distribution. The notable pattern from these figures is the widening racial differences in the test scores across the grades. African American boys and girls score the lowest among the three groups at all grades. Hispanic children, especially girls, are able to move faster than their African American counterparts beyond the first grade, although their initial scores are relatively low. Overall, girls do better than boys in reading test scores and Whites maintain their advantages in grade progression.

The key socioeconomic variables are the education and occupational prestige scores of the fathers and mothers. First, we convert the educational level into a continuous variable measured by years of schooling[7]. Considering the additional information obtained through the second round of parental interviews, and changes in educational level at the subsequent four waves, we are able to recover 695 cases for mothers' education and 1098 cases for fathers' education. Second, to reduce the missing values for the parental occupational prestige scores, we impute the values based upon the additional education information. We assign the average prestige scores in a particular level of education to those parents with missing values of prestige scores, and assign prestige scores to women not in the labor force with a similar logic[8]. Our imputation procedures push the mean prestige scores 5 points higher for mothers and 4 points higher for fathers. The overall higher prestige scores obtained through this imputation may result in part from the fact that some parents with lower education may actually be unemployed and out of the labor force.

---

[7] 8th grade or below=8, 9th-12th grade=10, high school diploma/equivalent=12, voc/tech program=13, Some college=14, Bachelor's degree=16, Graduate/professional school-no degree=17, Master's degree (MA, MS) =18, Doctorate or professional degree=21

[8] 10% of mothers who are out of the labor force have some college or a bachelor's degree.

## Methods

Our analysis has two stages. In the first stage, we search for a model to obtain a better estimate of the association between family background and children's achievement scores. We start with a recursive model and allow both parental education and occupation to affect the test scores of five waves. Our second step is to fit a Multiple Indicator Multiple Cause (MIMIC) model, which treats family background as a latent construct generated by parental education and occupation. This step enables the model to consider the measurement errors in parental characteristics and to predict the effect of the "unobserved" family background on children's test scores. In our third step, we consider five underlying, latent constructs corresponding to five waves of test scores, and add them to our step-two MIMIC model. This strategy allows us to evaluate the measurement errors in the test. In our final step, we add the causal chain between grades to the MIMIC model in the previous step. The causal chains derivate from Philips's (2000) theoretical arguments. Philips (2000) finds that year-to-year correlation and gain-initial score correlations are salient in elementary school, although smaller than in high school. These correlations suggest that children follow a trajectory of academic growth as they move through the grades, and their later rounds of test scores are determined by their adjacent previous achievement. Finally, we evaluate the goodness fits to obtain a better model, with chi-square value, BIC (the Bayesian Information Criteria) and RMSEA. A smaller chi-square value and a negative BIC indicate a better model fit.

After we find the best fit model for the whole sample, we proceed to compare group differences in the association between family background and children's test scores. We begin with separate modeling of gender-race specific samples. We then compare the gender and racial differences in three aspects: the measurements of family background by parental education and occupation, the measurements of test scores across five waves, the predicted effects of family background on test scores, and the causal chains between grades.  Similarly, we use chi-square value, BIC and RMSEA to find the best fitting model.

## Results

The upper panel of Table 1 shows the fits of the four models in our first stage of the analysis. Both the chi-square value and the BIC indicate that Model 4 is the best-fit model to estimate the association between family background and children's academic achievement. M1 is a recursive model and fits the data badly, with a chi-square value of 26846.86, and a BIC of 26759.5. After we treat family background as a latent construct, M2 significantly improves the model fit by a two-thirds reduction in the chi-square and the BIC values, with 12 more degrees of freedom. Although the measurement errors in M3 also improve the model fit by reducing the BIC from 8546.91 to 6691.82,this model still fits the data poorly. Finally, when we put in casual chains to estimate the grade-grade correlation in the test scores, the model fits the data nicely with a chi-square value of 33.51 and a negative BIC (-98.96). By this criterion, M4 with causal chains fits particularly well in the Black and Hispanic samples.

The model fit supports our speculation that it is necessary to consider the following components in estimating the association between family background and children's academic growth: a) allow the measurement errors in parental characteristics to correlate with each other; b) treat the family background as a latent construct; c) correct for measurement errors in the test scores, and d) build causal chains between grades as children move through the grades.

During the second stage of our analysis, we compare the group differences in the effect of causal chains and family background along gender and racial lines. In the analysis, we put equality constraints on the effects of parental characteristics, the latent family background and the latent test scores of the previous grade. These equality constraints mean that we force these estimated effects for the NH Blacks and Hispanics to be the same as those for the Whites. The lower panel of Table 1 shows the model fits by constraining the slopes. Since the model without constraints on the slopes and errors in the latent variables fits better, we conclude that there are gender-race specific slopes of family background and causal chains.

Table 2 shows the race specific estimated effects of parental education and occupational prestige scores on the latent family background for boys. The standardized coefficients of parental education and occupational prestige scores show that Hispanic children share about the same pattern as Whites, suggesting that parental education is more important than occupation in determining the latent family background. But, among African

American children, both parental occupation and education matter to measure the latent family background, and the effect of parental occupational prestige scores is more than three times larger than that of Whites and Hispanics (0.42 vs. 0.12 vs. 0.13). Furthermore, the raw coefficient of parental occupational prestige scores seems to insignificant for Whites and Blacks, suggesting that current measurement of prestige scores is too crude to capture the effect.

Table 3 shows the estimated slopes of the observed test scores on the latent test scores. The racial differences in kindergarten are smaller and the observed test scores of Hispanic children are subject to fewer errors than those of Whites and Blacks. But the racial differences reverse direction starting from the spring semester of the first grade. The measurement errors for Hispanic children are consistently larger than those for Whites, and the test scores of Black students do not necessarily contain similar problems. This set of evidence indicates that the measurement errors perhaps do not occur randomly for Hispanic and Black students. As children move through the grades, potential “content bias” may exert a large effect on the measures of the reading skills for children whose primary home language is not English.

Table 4 reports the estimated slopes of the latent family background on the latent test scores. The common patterns shared by the three racial groups are that the effects of the latent family background get stronger after the first grade and the effect sizes for all three racial groups almost double in the third and fifth grades. The academic achievement of African American children is under the strongest influence of family background from kindergarten to third grade, while White and Hispanic children share a similar size of influence across the grades. By the fifth grade, the effects of family for White and Hispanic children catch up with those for African American children.

Table 5 shows the estimated slopes of the causal chains in the test scores between the grades. All the coefficients of the causal chains are statistically significant for all groups, indicating that children's growth in reading skills shows a "path dependent" pattern. Children's kindergarten performance strongly affects their reading IRT scores in the first grade, yet the effect of test scores in the first and third grade is slightly weakened. The causal effect of previous scores on subsequent test scores supports the notion that children's learning process is cumulative and the correlation is larger than Philips (2000) argues. Recall that family background's influence doubles in the third and fifth grade. It is reasonable to conclude that the transmission of social inequality has started earlier and made a difference in children's initial education career. Children who start with lower scores may find themselves falling behind and may find it hard to catch up if those with good scores maintain their "momentum" of advancement in subsequent school years. But, the racial differences in the causal chains are small. For the White and African American children, the reading scores at kindergarten have a larger effect, while Hispanic children are subject to a stronger effect of the reading scores of the first grade.

In summary, we find that parental education and occupational prestige scores are stronger predictors of family background for African American children and the effect of family background becomes larger in the third and fifth grade. The measurement errors in the test scores are modest for Hispanic children starting from the first grade and all children obtain their learning gains accumulatively, heavily relying on their previous test scores.

## Conclusions

In this analysis, we examine the association between family background and children's reading test scores in the elementary school years. With the structural equation modeling approach, we make our contribution in the following ways: first, we are able to show that a regular regression model (recursive model) is likely to introduce bias into the estimates without considering the measurement errors in parental characteristics and children's test scores. Second, the MIMIC model with causal chains provides a nice framework and model to examine the cumulative process of children's learning, with correction of measurement errors. Thirdly, we find that family background, even simply measured by parental education and crude occupational prestige scores, shows a strong influence on children's academic achievement in the elementary school years. The consistent and strong causal chains indicate that the racial test scores gap is likely to widen if some of the minority students with lower initial scores are not adequately remedied.  In addition, our finding suggests that part of the racial test scores gap may result from the less precise measures of Hispanic children's test scores starting from the first grade.

However, our analysis is not immune to limitations. First, the Black sample is small and highly selected and thus may not represent the test scores growth for their group. Second, we only analyze children who consistently remain in the study from kindergarten to fifth grade, while leaving attrition unaddressed. It is possible that higher rates of attrition among African American children contribute partly to the small numbers of the Black sample. Third, we notice that minority children suffer from a high frequency of school movement and grade retention (Hauser 2004). So, the roles of moving and retention should be examined in greater detail in future research. Fourth, our measures of parental characteristics are limited to education and occupational prestige scores, and the prestige scores are very coarsely measured in the ECLS-K. The future study should use refined measures of occupation and include more important features of family background, such as household income and family wealth.

## References


1. Bielby, William T., and Hauser, Robert M. 1977 "Response Error in Earnings Functions for Nonblack Males," *Sociological Methods and Research*. 6. pg.241-280.
2. Downey, Douglas B; Paul T von Hippel; Beckett A Broh. 2004. Are Schools the Great Equalizer? Cognitive Inequality during the Summer Month. *American Sociological Review*; 69, 5; pg. 613.
3. Fryer Roland G. Jr. and Steven D. Levit, 2004. Understanding the Black-White Test Score Gap in the First Two Years of School. *The Review of Economics and Statistics*; Volume. LXXXVI: 2. pg. 447
4. Grissmer, David W., Sheila N. Kirby, Mark Berends, and Stephanie Williamson. 1994. Student Achievement and the Changing American Family. Washington, D.C.: RAND Institute on Education and Training.
5. Hauser, Robert M. 2004. "Schooling and Academic Achievement in Time and Place." *Research in Social Stratification and Mobility*, 21, edited by K. Leicht. London: Elsevier. pg. 47
6. Hauser, Robert M. and Megan Andrew. 2007. Reliability of Student and Parent Reports of Socioeconomic Status in NELS-88, working paper, presented at ITP seminar at UW-Madison, http://www.wcer.wisc.edu/itp/Spring%2008%20seminar/HauserNELS-SES%20measurement_070607a.pdf.
7. Jencks, Christopher and Meredith Phillips. 1998. The black-white test score gap. Washington, D.C.: Brookings Institution Press.
8. Jencks, Christopher. 1998, Racial Bias in Testing, Pp. 55-85 in Jencks, Christopher and Meredith Phillips. 1998. The black-white test score gap. Washington, D.C.: Brookings Institution Press.
9. Kao, Grace and Jennifer S Thompson, 2003 Racial and ethnic stratification in educational achievement and attainment, *Annual Review of Sociology*; 29, pg. 417
10. Phillips, Meredith. 2000. "How to Improve Our Understanding of Ethnic Differences in Academic Achievement: Empirical Lessons from National Data" Pp. 103-132 in Analytic Issues in the Assessment of Student Achievement, edited by D. W. Grissmer and J. M. Ross. Washington, DC: National Center for Education Statistics.
11. Ream, Robert K, 2005. Toward Understanding How Social Capital Mediates the Impact of Mobility on Mexican American achievement. *Social Forces*. 84, 1; pg. 201
12. Yeung, W. Jean and Dalton Conley, 2008. Black – White Achievement Gap and Family Wealth, *Child Development*,79: 2, pg. 303.

**Table1. Goodness Fit of selected models**

| Models | Specification | N | Chi-square | BIC | df | RMSEA |
|---|---|---|---|---|---|---|
| M1 | Recursive path model | 6846 | 26846.86 | 26758.55 | 10 | 0.626 |
| M2 | MIMIC w/ latent family background | 6846 | 8758.86 | 8546.91 | 24 | 0.231 |
| M3 | M2 + latent test scores model | 6846 | 6877.28 | 6691.82 | 21 | 0.218 |
| M4 | M3 + causal chains | 6846 | **33.51** | **-98.96** | 15 | 0.013 |
| | NH White | 5221 | 28.34 | -100.07 | 15 | 0.013 |
| | NH Black | 584 | 15.60 | -79.95 | 15 | 0.008 |
| | Hispanic | 1041 | 22.36 | -81.86 | 15 | 0.022 |
| | Male | 3453 | 25.04 | -97.16 | 15 | 0.014 |
| | Female | 3393 | 36.83 | -85.11 | 15 | 0.021 |
| **Male** | | | | | | |
| M5 | Everything is equal | 3453 | 110.99 | -581.50 | 85 | 0.016 |
| M6 | Free betas | 3453 | 100.07 | -396.90 | 69 | 0.02 |
| M7 | Free gammas | 3453 | 100.06 | -527.26 | 77 | 0.016 |
| M8 | Free betas & gammas | 3453 | 87.97 | -409.00 | 61 | 0.02 |
| M9 | Free betas & gammas & TE & PS | 3453 | 57.85 | -308.76 | 45 | 0.016 |
| **Female** | | | | | | |
| M10 | Everything is equal | 3393 | 135.69 | -555.31 | 85 | 0.023 |
| M11 | Free betas | 3393 | 125.54 | -435.39 | 69 | 0.027 |
| M12 | Free gammas | 3393 | 124.46 | -501.51 | 77 | 0.023 |
| M13 | Free betas & gammas | 3393 | 112.78 | -383.12 | 61 | 0.027 |
| M14 | Free betas & gammas & TE & PS | 3393 | 66.83 | -396.90 | 45 | 0.021 |

**Table2. Estimated slopes of parental education and occupation on family background for boys**

| | NH White | NH Black | Hispanic |
|---|---|---|---|
| Mother's education | 0.50<br>(0.00) | 0.24<br>(0.00) | 0.57<br>(0.00) |
| Mother's occupational prestige score | 0.10<br>(0.01) | 0.42<br>(0.01) | 0.02<br>(0.01) |
| Father's education | 0.48<br>(0.00) | 0.31<br>(0.00) | 0.45<br>(0.00) |
| Father's occupational prestige score | 0.12<br>(0.01) | 0.42<br>(0.01) | 0.13<br>(0.01) |

Note: standardized coefficients.

**Table3. Estimated slopes of latent test scores on observed test scores for boys**

| | NH White | NH Black | Hispanic |
|---|---|---|---|
| Reading IRT scale scores at Fall of Kindergarten | 0.96<br>[0.08] | 0.95<br>[0.10] | 0.98<br>[0.04] |
| Reading IRT scale scores at Spring of Kindergarten | 0.97<br>[0.06] | 0.97<br>[0.06] | 0.99<br>[0.02] |
| Reading IRT scale scores at Spring of first grade | 0.99<br>[0.02] | 1.00<br>[0.00] | 0.94<br>[0.12] |
| Reading IRT scale scores at Spring of third grade | 0.95<br>[0.10] | 0.96<br>[0.08] | 0.94<br>[0.12] |
| Reading IRT scale scores at Spring of fifth grade | 0.95<br>[0.10] | 0.96<br>[0.08] | 0.94<br>[0.12] |

Note: standardized coefficients, the error variances of the estimate are in [ ]. For example, for white children at fall of kindergarten, $e = 1 - 0.96^2 = 0.08$

**Table4. Estimated slopes of latent family background on latent test scores**

| | NH White | | NH Black | | Hispanic | |
|---|---|---|---|---|---|---|
| | raw coef. | std. coef | raw coef. | std. coef | raw coef. | std. coef |
| Reading IRT scale scores at Fall of Kindergarten | 1.00 | 0.28 | 1.00 | 0.30 | 1.00 | 0.27 |
| Reading IRT scale scores at Spring of Kindergarten | 1.15 | 0.32 | 1.42 | 0.43 | 1.12 | 0.30 |
| | (0.04) | | (0.04) | | (0.04) | |
| Reading IRT scale scores at Spring of first grade | 1.71 | 0.48 | 2.08 | 0.63 | 1.77 | 0.47 |
| | (0.09) | | (0.09) | | (0.09) | |
| Reading IRT scale scores at Spring of third grade | 2.60 | 0.74 | 2.77 | 0.84 | 2.47 | 0.66 |
| | (0.16) | | (0.16) | | (0.16) | |
| Reading IRT scale scores at Spring of fifth grade | 2.79 | 0.79 | 2.57 | 0.78 | 2.60 | 0.69 |
| | (0.17) | | (0.17) | | (0.17) | |

**Table5. Estimated slopes of causal chains between latent test scores**

| | **NH White** | | **NH Black** | | **Hispanic** | |
|---|---|---|---|---|---|---|
| | raw coef. | std. coef | raw coef. | std. coef | raw coef. | std. coef |
| Reading IRT scores at Fall of Kindergarten | 1.17 | 0.90 | 1.09 | 0.83 | 1.06 | 0.82 |
| | (0.03) | | (0.03) | | (0.03) | |
| Reading IRT scores at Spring of Kindergarten | 1.10 | 0.75 | 1.19 | 0.79 | 1.09 | 0.79 |
| | (0.02) | | (0.02) | | (0.02) | |
| Reading IRT scores at Spring of first grade | 0.86 | 0.71 | 0.87 | 0.71 | 0.97 | 0.76 |
| | (0.02) | | (0.02) | | (0.02) | |
| Reading IRT scores at Spring of third grade | 0.90 | 0.90 | 0.91 | 0.92 | 0.90 | 0.91 |
| | (0.02) | | (0.02) | | (0.02) | |

**Table6. Reduced forms of parental characteristics on children's test scores**

| | NH Whites | | | | NH Blacks | | | | Hispanics | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MOMED | MOMOC | DADED | DADOC | MOMED | MOMOC | DADED | DADOC | MOMED | MOMOC | DADED | DADOC |
| ETA 1 | 0.03 | 0.02 | 0.03 | 0.03 | 0.02 | 0.11 | 0.02 | 0.11 | 0.03 | 0.00 | 0.03 | 0.03 |
| | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) |
| ETA 2 | 0.03 | 0.02 | 0.03 | 0.03 | 0.02 | 0.11 | 0.02 | 0.11 | 0.03 | 0.00 | 0.03 | 0.03 |
| | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) | 0.00 | (0.01) |
| ETA 3 | 0.04 | 0.03 | 0.03 | 0.03 | 0.02 | 0.16 | 0.03 | 0.16 | 0.04 | 0.01 | 0.03 | 0.03 |
| | 0.00 | (0.02) | 0.00 | (0.02) | (0.01) | (0.02) | (0.01) | (0.02) | 0.00 | (0.02) | 0.00 | (0.02) |
| ETA 4 | 0.05 | 0.04 | 0.05 | 0.05 | 0.03 | 0.23 | 0.04 | 0.24 | 0.06 | 0.01 | 0.04 | 0.05 |
| | (0.01) | (0.02) | (0.01) | (0.02) | (0.01) | (0.03) | (0.01) | (0.03) | (0.01) | (0.02) | (0.01) | (0.03) |
| ETA 5 | 0.08 | 0.06 | 0.08 | 0.08 | 0.04 | 0.31 | 0.06 | 0.32 | 0.08 | 0.01 | 0.06 | 0.07 |
| | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.03) | (0.01) | (0.03) |
| ETA 6 | 0.09 | 0.07 | 0.08 | 0.08 | 0.04 | 0.28 | 0.05 | 0.29 | 0.09 | 0.01 | 0.07 | 0.08 |
| | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.04) | (0.01) | (0.04) |

Figure 1. Model 1-Recursive model

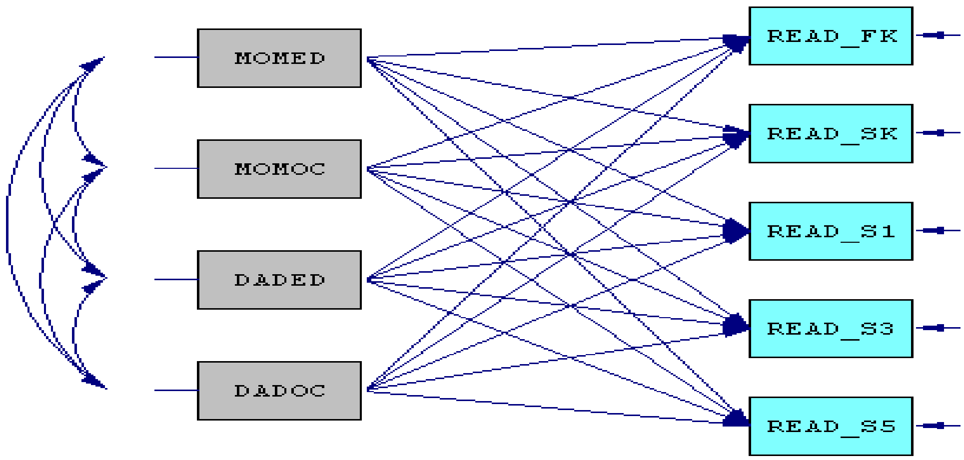



Figure 2. Model 2-MIMIC model with latent family background

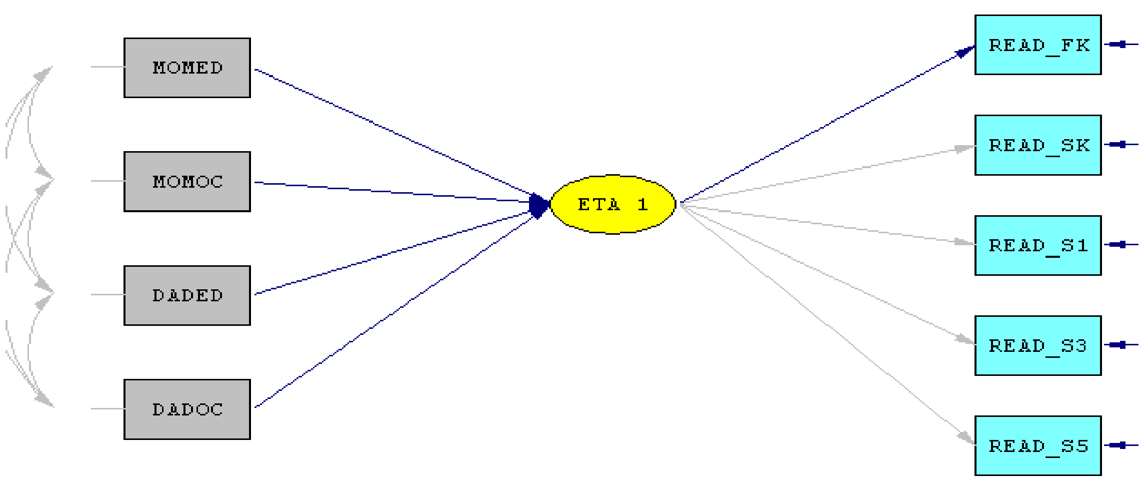


Figure 3. Model 3- MIMIC model with family background and latent test scores

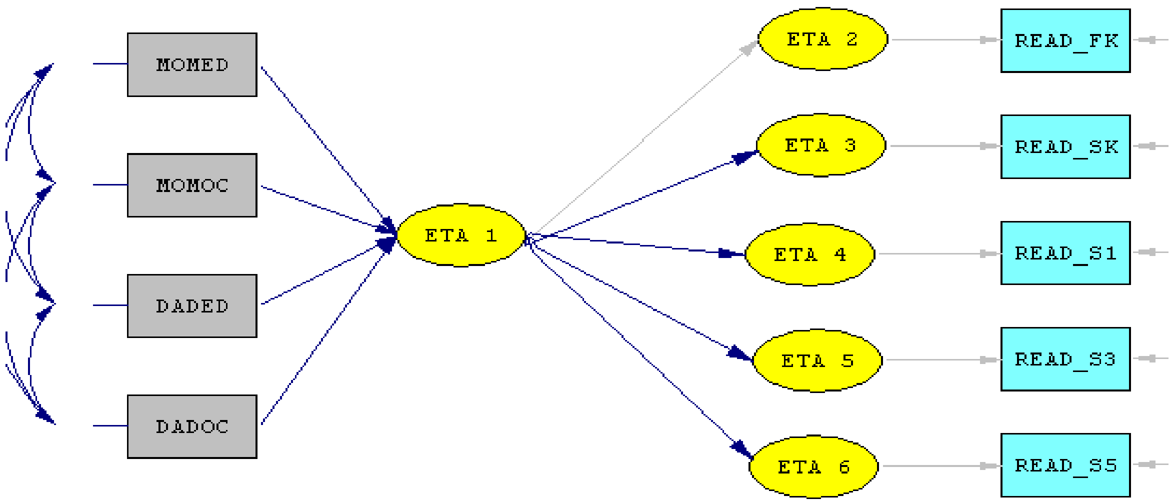



Figure 4. Model 4- MIMIC model with causal chains

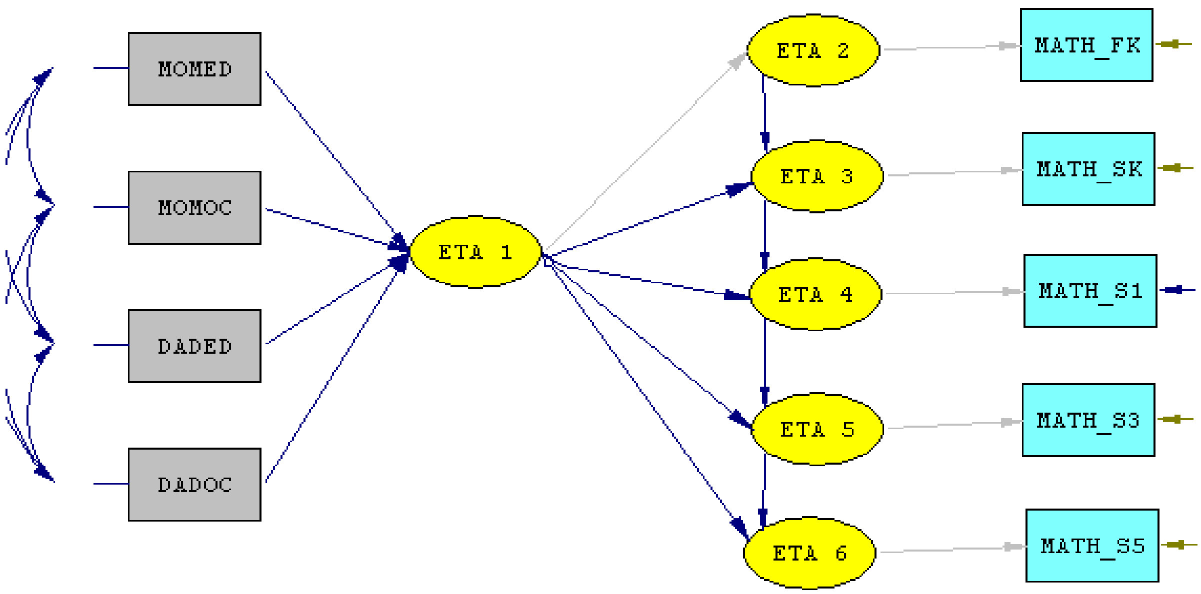


Figure 5. Standardized Estimates, Model 4 Female NH White

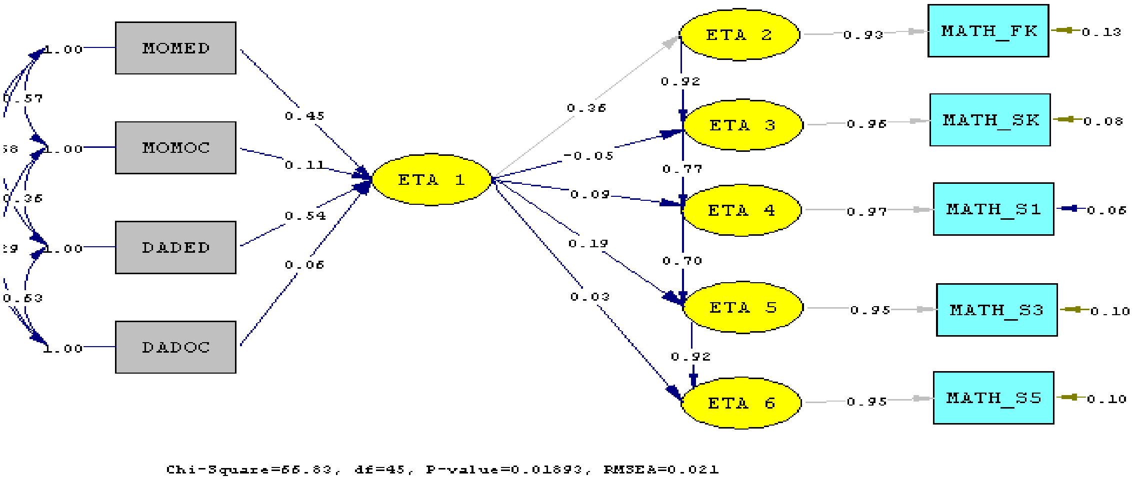



Figure 6. Standardized Estimates, Model 4 Female NH Blacks

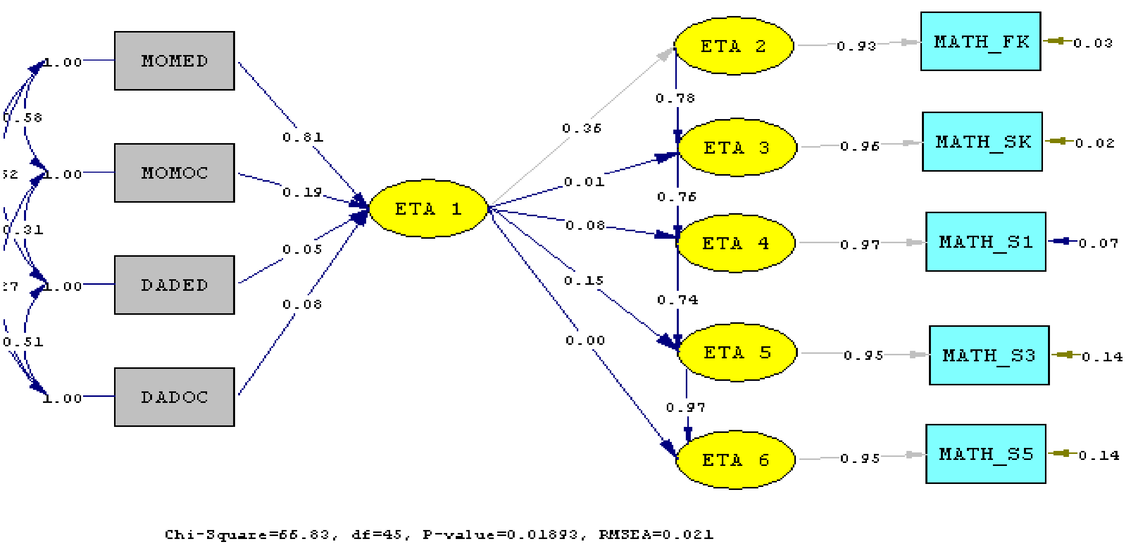


Figure 7.  Standardized Estimates, Model 4   Female Hispanics

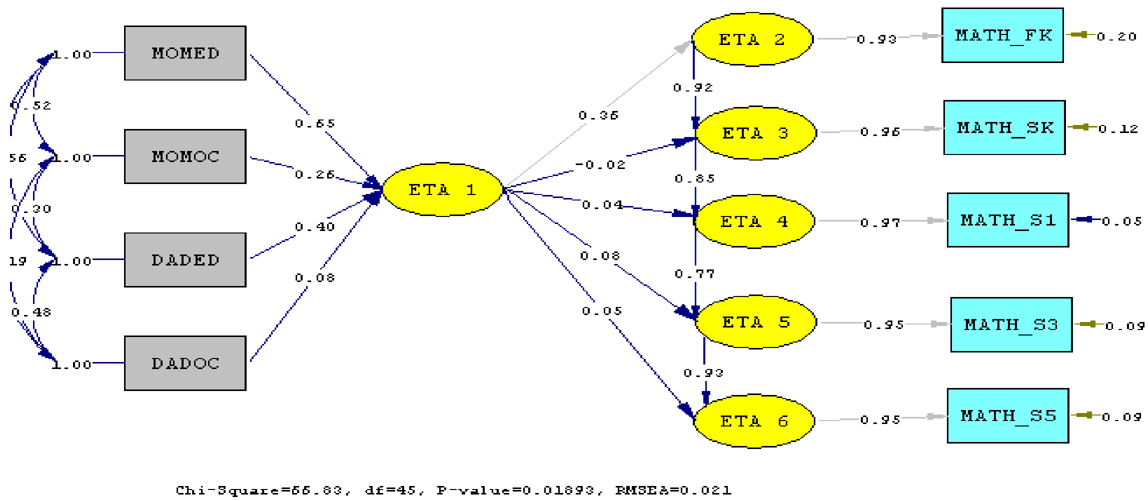


Figure 8. Standardized Estimates, Model 4 Male NH White

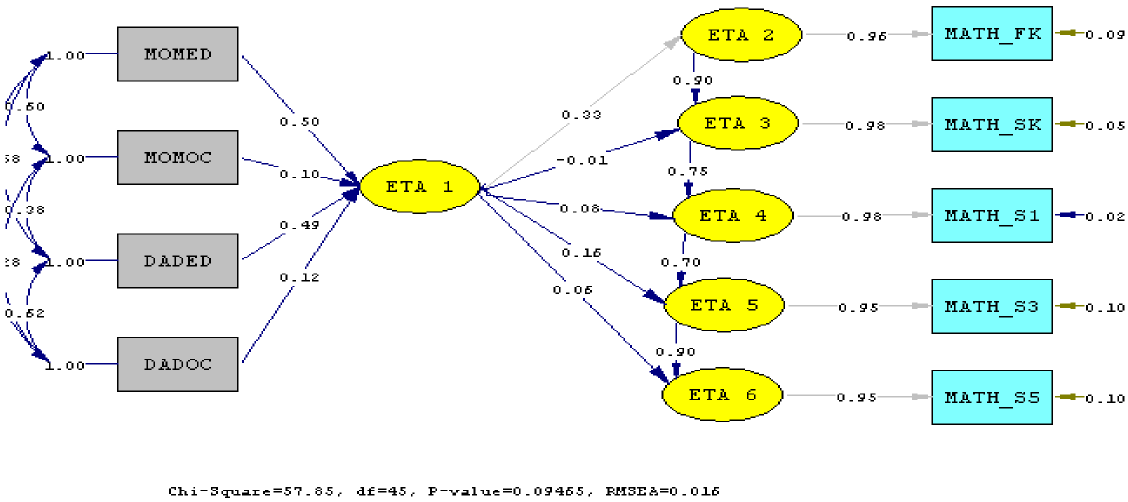



Figure 9. Standardized Estimates, Model 4 Male NH Blacks

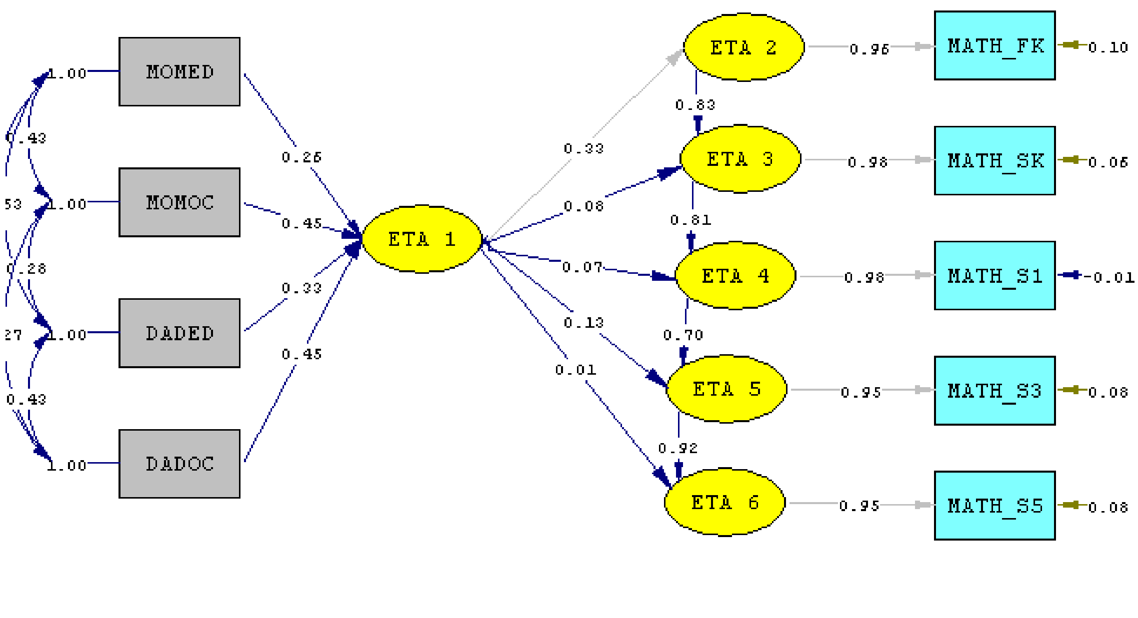
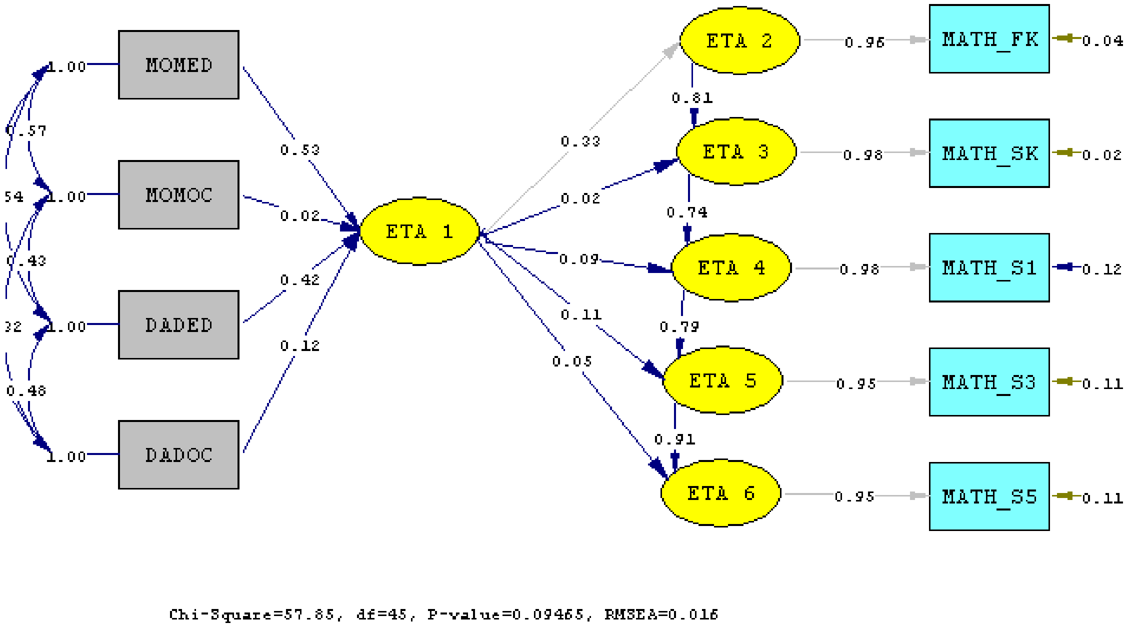


Figure 10.  Standardized Estimates, Model 4   Male Hispanics



Appendix1

Figure 1. IRT Scale Score at Fall of Kindergarten

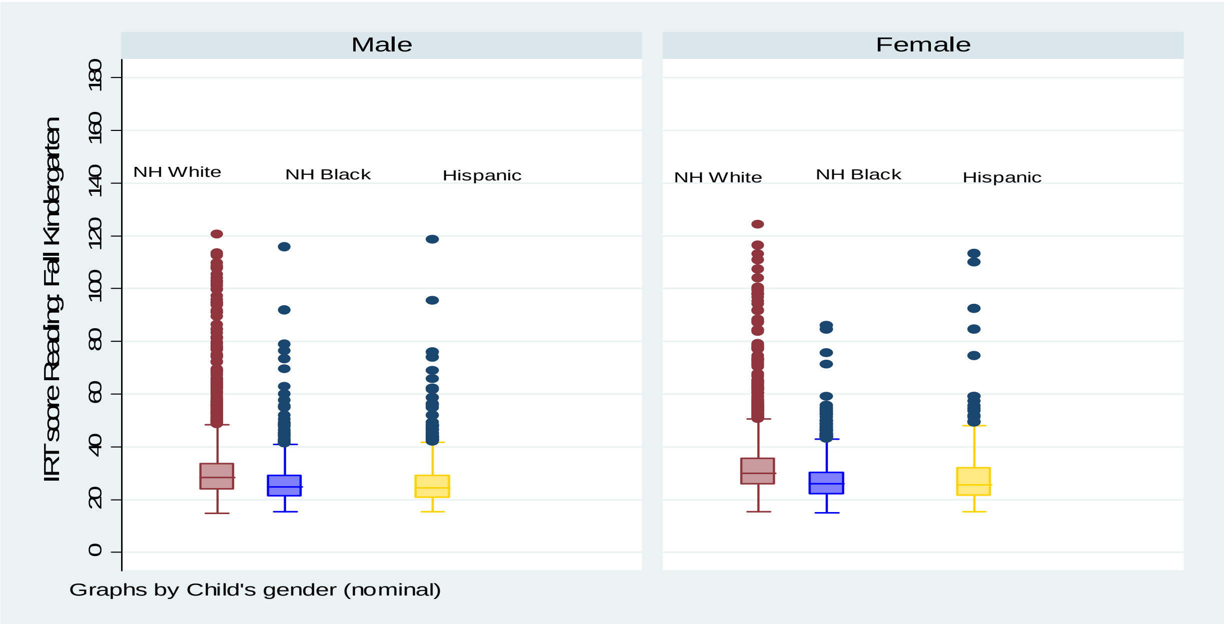



Figure 2. IRT Scale Score at Spring of Kindergarten

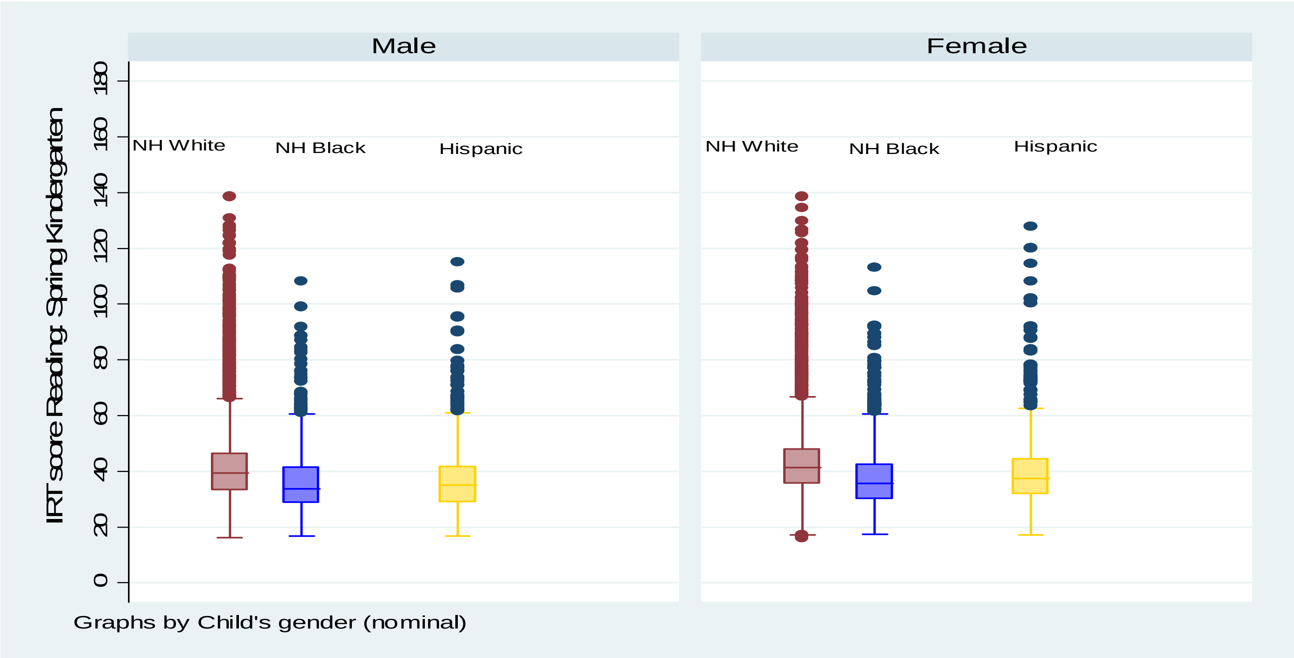


Figure 3. IRT Scale Score at Spring of First Grade

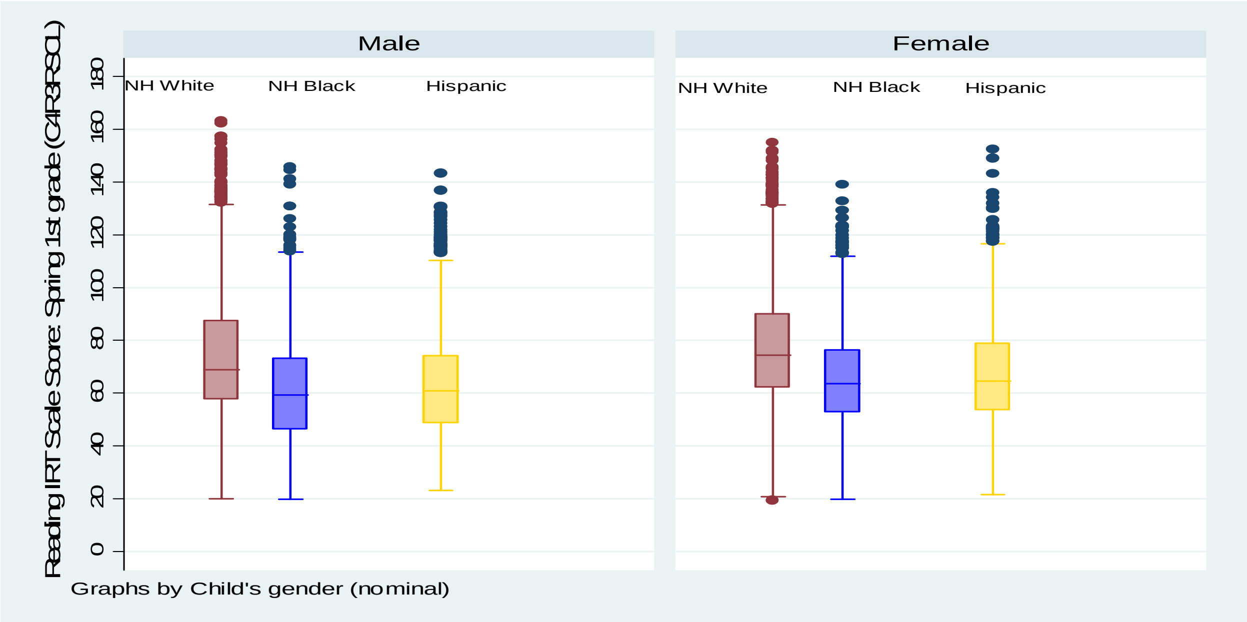



Figure 4. IRT Scale Score at Spring of Third Grade

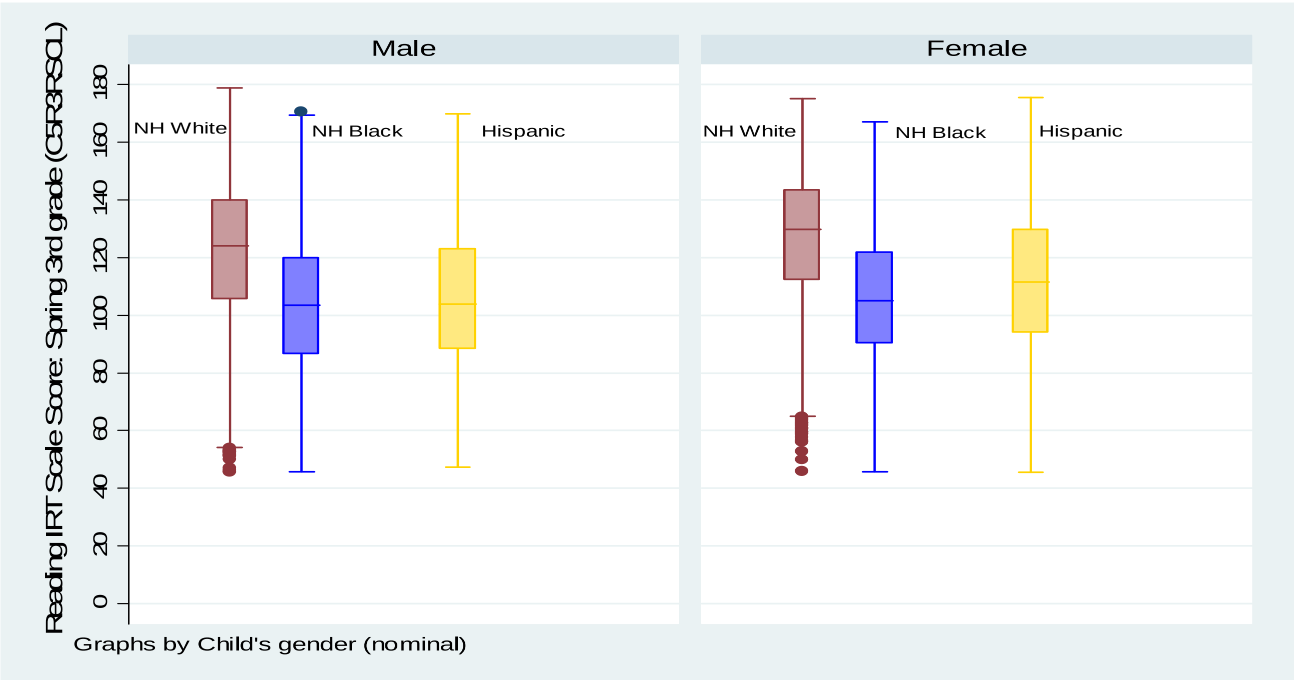


Figure 5.  IRT Scale Score at Spring of Fifth Grade