

Understanding the Black-White Test Score Gap in the First Two Years of School

R. Fryer & S. Levitt

Presented by Owen Zidar

February 6, 2012

- **Question:** How big is the black-white test score gap in the first two years of school and why does it exist?
- **Data:** Test scores for 20K+ kindergarteners in 1998 from Dept. of Ed
- **Results:**
 - Black kindergartners score $.64\sigma$ worse than whites
 - Surprisingly, in this data, controlling for a few things eliminates the gap
 - BUT blacks lose substantial ground over first two years of school
 - Suggestive evidence these losses due to school quality differences

Why study the black-white test score gap?

- **Huge labor market impacts:** Most of the observed black-white wage differentials among adults disappear once lower 8th grade test scores are taken into account
- **Prior Literature:** Previous attempts to explain the gap have fallen short - a large residual remained for black students even after controlling for many covariates
- **Policy Implications:** Knowing the sources of divergence may aid in developing policies to alleviate the problem

• Data:

- Early Childhood Longitudinal Study kindergarten cohort (ECLS-K) is a nationally representative sample of 20K children entering kindergarten in 1998
- Interviewed at the fall and spring of kindergarten and the spring of first grade
- LHS variables: math and reading test scores (and teacher assessments of math and reading skills)
- RHS variables: Race, SES, gender, child's age at enrollment, WIC participation ¹, mother's age at first birth, birthweight and number of kids books at home.

• Regression:

$$TESTSCORE_i = RACE_i' \gamma + X_i' \beta + \epsilon_i$$

¹WIC is a nutrition program aimed at low income mothers and children

Gap with and without Controls

TABLE 2.—THE ESTIMATED BLACK-WHITE TEST SCORE GAP IN FALL OF KINDERGARTEN

Variables	Math					Reading				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Black	-.638 (.022)	-.368 (.022)	-.238 (.023)	-.094 (.023)	-.102 (.026)	-.401 (.024)	-.134 (.025)	-.006 (.026)	.117 (.025)	.093 (.030)
Hispanic	-.722 (.022)	-.429 (.023)	-.302 (.022)	-.203 (.022)	-.171 (.028)	-.427 (.027)	-.223 (.026)	-.137 (.026)	-.064 (.025)	-.076 (.029)
Asian	.150 (.056)	.070 (.051)	.190 (.051)	.265 (.048)	.274 (.050)	.335 (.064)	.256 (.059)	.371 (.059)	.409 (.058)	.375 (.060)
Other race	-.503 (.041)	-.329 (.037)	-.253 (.036)	-.158 (.035)	-.113 (.035)	-.401 (.044)	-.230 (.040)	-.155 (.040)	-.072 (.038)	-.014 (.039)
Socioeconomic status composite measure	—	.456 (.014)	.389 (.014)	.302 (.014)	.072 (.024)	—	.451 (.014)	.393 (.015)	.299 (.015)	.092 (.023)
Number of children's books	—	—	.007 (.001)	.006 (.001)	.005 (.001)	—	—	.007 (.001)	.006 (.001)	.004 (.001)
(Number of children's books) ² (×1000)	—	—	-.023 (.003)	-.020 (.002)	-.027 (.016)	—	—	-.025 (.003)	-.021 (.003)	-.017 (.017)
Female	—	—	—	.010 (.015)	.00 (.015)	—	—	.159 (.017)	.153 (.017)	.153 (.016)
Age at kindergarten fall (in months)	—	—	—	.056 (.002)	-2.680 (.542)	—	—	.041 (.002)	-2.409 (.483)	-2.409 (.483)
Birth weight (ounces) (× 10)	—	—	—	.029 (.004)	.030 (.004)	—	—	.019 (.004)	.022 (.004)	.022 (.004)
Teenage mother at time of first birth	—	—	—	-.109 (.018)	-.029 (.021)	—	—	-.144 (.020)	-.069 (.022)	-.069 (.022)
Mother at least 30 at time of first birth	—	—	—	.182 (.025)	.11 (.028)	—	—	.226 (.027)	.155 (.030)	.155 (.030)
WIC participant	—	—	—	-.211 (.019)	-.120 (.020)	—	—	-.184 (.021)	-.104 (.021)	-.104 (.021)
R ²	0.108	0.223	0.239	0.317	0.354	0.045	0.16	0.175	0.233	0.279
Number of observations	13,290					12,601				
Full set of covariates included in regression?	N	N	N	N	Y	N	N	N	N	Y

NOTES: The dependent variable is the math or reading test score in the fall of kindergarten. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic whites are the omitted race category, so all of the race coefficients are gaps relative to that group. The unit of observation in a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set. In addition to the variables included in the table, indicator variables for students with missing values on each covariate are also included in the regressions. In addition, columns 5 and 10 report only a subset of the coefficients from regressions with 98 covariates included in the specification. The full results for columns 5 and 10 are reported in appendix table A1. Note that the specifications in columns 5 and 10 include age and age squared; that is why the coefficient on age changes so dramatically relative to other columns in the table.

They look at subsamples to check the robustness of their results

- **Common Support:** Could be concerned that restricting coefficients to be identical across sample causes misleading results, so they run regressions on common support (e.g. only single mothers in one region of the country)
- **Equal Responsiveness:** If black children do not derive as much benefit from improvements in SES, number of books, etc, then earlier results will overstate convergence in test scores

Falling Behind: Evolution of Gap

TABLE 5.—THE EVOLUTION OF TEST SCORE GAPS BY RACE AS CHILDREN AGE

Variable	Math			Reading		
	Fall Kindergarten	Spring Kindergarten	Spring First Grade	Fall Kindergarten	Spring Kindergarten	Spring First Grade
Black	-.094 (.023)	-.201 (.025)	-.250 (.028)	.117 (.025)	-.009 (.027)	-.071 (.029)
Hispanic	-.203 (.022)	-.187 (.024)	-.120 (.026)	-.064 (.025)	-.005 (.027)	.001 (.029)
Asian	.265 (.048)	.221 (.049)	.115 (.044)	.409 (.058)	.434 (.054)	.345 (.045)
Other race	-.158 (.035)	-.166 (.039)	-.195 (.042)	-.072 (.038)	-.099 (.039)	-.163 (.042)
SES composite measure	.302 (.014)	.284 (.014)	.263 (.014)	.299 (.015)	.280 (.015)	.284 (.014)
Number of books	.006 (.001)	.006 (.001)	.005 (.001)	.006 (.001)	.005 (.001)	.006 (.001)
(Number of books) ² × 1000	.020 (.002)	-.019 (.003)	-.019 (.003)	-.021 (.003)	-.020 (.003)	-.022 (.003)
Female	.010 (.015)	.003 (.016)	-.033 (.017)	.159 (.017)	.195 (.017)	.216 (.017)
Age at kindergarten fall (in months)	.056 (.002)	.051 (.002)	.036 (.002)	.041 (.002)	.034 (.002)	.021 (.002)
Birth weight (ounces) × 100	.029 (.004)	.003 (.000)	.029 (.004)	.019 (.004)	.002 (.000)	.024 (.005)
Teenage mother at time of first birth	-.109 (.018)	-.112 (.021)	-.111 (.022)	-.144 (.020)	-.138 (.021)	-.131 (.024)
Mother in 30s at time of first birth	.182 (.025)	.127 (.024)	.093 (.022)	.226 (.027)	.158 (.025)	.085 (.024)
WIC participant	-.211 (.019)	-.195 (.020)	-.201 (.021)	-.184 (.021)	-.152 (.02)	-.182 (.022)
R ²	0.317	0.282	0.240	0.233	0.197	0.194
Number of obs.	13,290	13,290	13,290	12,601	12,601	12,601

NOTES: The dependent variable is fall kindergarten test scores in columns 1 and 3 and spring first grade test scores in columns 2 and 4. All specifications include the parsimonious set of controls corresponding to columns 4 and 9 of table 2. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic whites are the omitted race category, so all of the race coefficients are gaps relative to that group. The unit of observation is a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set. In addition to the variables included in the table, indicator variables for students with missing values on each covariate are also included in the regressions.

Why are blacks falling behind in first two years?

- 1 Attend lower quality schools on average
- 2 Importance of parental & environmental contributions grow over time
- 3 Summer setbacks
- 4 Artifact of testing or poor measurement, rather than true losses
- 5 Interaction between black students and schools interferes with learning process (e.g. discrimination or low expectations, systematic differences in self-control or socialization, etc)
- 6 The fall kindergarten test scores are measuring a different set of skills than the later tests, and the gap between whites and blacks is greater on the set of skills measured later

Evaluating the Explanations (1 of 2)

Some suggestive evidence for School Quality Differences

- ① **Little Overlap:** One third of the 1,000 schools in the sample have no black students. The mean black student attends a school that is 59% black and 8% Hispanic, whereas the mean white student's school is 6% black and 5% Hispanic
- ② **School Fixed Effects:**² Estimates from comparing black and white students at the same school shrink the gap to one-third of the original magnitude.³
- ③ **Observable Differences:** While traditional measures of quality are similar, non-traditional measures (e.g. % of free lunch students, degree of gang problems and non-student loitering) are worse

²To the extent that differential average school quality across races completely explains the gap, one would predict the gap shouldn't widen overtime when comparing blacks and whites at the same school

³Concern that whites who go to school with blacks have lower trajectories is alleviated by evidence that these whites do better on kindergarten exams

Other explanations less convincing

- ① *Importance of parental & environmental contributions grow over time:* **No.** SES and other coefficients smaller for 1st grade unlike school safety measures. Residual gap after controlling for environment grows.
- ② *Summer setbacks:* **No.** Little Δ in pre- and post- summer scores
- ③ *Artifact of testing:* **No.** Subjective teacher measures show similar gap
- ④ *Teacher Bias:* **No.** Students with black teachers have similar outcomes
- ⑤ *Kindergarten tests evaluate different skills:* **No, but interesting** B & W students do similarly on counting, numbers and shapes, but worse on multiplication division when entering kindergarten. Lost ground over two years is attributable to +, -, \times , and \div .

Summary of findings:

- 1 Black kindergartners score $.64\sigma$ worse than whites, but controlling for a few things eliminates this gap
- 2 Blacks lose substantial ground over first two years of school
- 3 Suggestive evidence these losses are due to school quality differences, but Asians and Hispanics are also attending worse schools and are not losing ground

APPENDIX: Summary stats LHS

TABLE 1.—SUMMARY STATISTICS BY RACE: STUDENT CHARACTERISTICS

Variable	Full Sample	White	Black	Hispanic	Asian
Test Scores:					
Fall kindergarten math	.026 (1.038)	.274 (1.073)	-.364 (.765)	-.448 (.898)	.424 (1.275)
Spring first grade math	.013 (1.092)	.249 (.984)	-.479 (.977)	-.314 (1.040)	.272 (1.089)
Fall kindergarten reading	.014 (1.010)	.147 (1.073)	-.255 (.892)	-.280 (.944)	.481 (1.459)
Spring first grade reading	.033 (1.123)	.175 (1.073)	-.354 (1.062)	-.134 (1.062)	.472 (1.088)
Subjective teacher assessments:					
Fall kindergarten math	.071 (1.115)	.249 (1.084)	-.178 (1.001)	-.286 (.988)	.129 (1.269)
Spring first grade math	-.005 (1.115)	.130 (1.084)	-.355 (1.101)	-.163 (1.102)	.234 (.983)
Fall kindergarten reading	.081 (1.041)	.267 (1.050)	-.114 (1.025)	-.354 (1.043)	.063 (1.212)
Spring first grade reading	.006 (1.145)	.118 (1.050)	-.252 (1.177)	-.146 (1.127)	.206 (1.095)

APPENDIX: Summary stats RHS

Race:					
White	.578 (.550)	1.000	.000	.000	.000
Black	.161 (.390)	.000	1.000	.000	.000
Hispanic	.189 (.390)	.000	.000	1.000	.000
Asian	.028 (.130)	.000	.000	.000	1.000
Other	.045 (.260)	.000	.000	.000	.000
Other controls:					
Female	.486 (.520)	.481 (.594)	.495 (.536)	.494 (.542)	.503 (.591)
Age (in months), fall of kindergarten	67.061 (4.811)	67.402 (4.752)	66.877 (4.776)	66.413 (4.773)	65.927 (4.639)
SES composite measure	-.018 (.910)	.202 (.792)	-.359 (.780)	-.423 (.759)	.312 (1.034)
Number of children's books in the home	71.946 (64.551)	93.121 (64.792)	39.014 (41.986)	40.849 (47.752)	49.333 (55.859)
Mother's age at time of first birth	23.342 (6.052)	24.579 (5.916)	20.548 (5.139)	21.762 (5.256)	25.807 (6.330)
Child's birth weight (in ounces)	117.827 (23.847)	120.256 (22.942)	110.315 (25.373)	117.091 (23.048)	112.667 (22.498)
WIC participation	.475 (.517)	.332 (.493)	.772 (.481)	.652 (.539)	.313 (.529)
Frequency of missing values:					
Missing number of children's books in the home	.010 (.130)	.010 (.099)	.008 (.097)	.007 (.108)	.017 (.148)
Missing WIC	.012 (.130)	.007 (.099)	.029 (.195)	.014 (.108)	.010 (.118)
Missing mother's age	.020 (.130)	.008 (.099)	.049 (.244)	.022 (.163)	.045 (.207)
Missing birth weight	.128 (.390)	.100 (.297)	.150 (.390)	.170 (.434)	.246 (.502)

NOTES: The entries are means and standard deviations of student-level data for those students in ECLS-K who do not have missing values for test scores, race, or age. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Subjective teacher assessments have also been normalized in this manner. The category white includes only non-Hispanic whites. Precise definitions of the variables are provided in the data appendix. The SES composite measure incorporates information on parental education, occupational status, and family income. The SES measure ranges from -4.75 to 2.75 in the sample, with larger numbers indicating higher socioeconomic status. The total number of students in the sample who receive a positive weight in the estimation is 13,290. The bottom panel of the table reports the frequency of missing values for the covariates. In all cases, sample weights provided with ECLS are used in the calculations.

APPENDIX: Different results, so check robustness

TABLE 3.—SENSITIVITY ANALYSIS AND EXTENSIONS OF THE BASIC MODEL FOR FALL KINDERGARTEN TEST SCORES

Specification	Coefficient on Black for:		Coefficient on Hispanic for:		Coefficient on Asian for:	
	Math	Reading	Math	Reading	Math	Reading
Baseline	-.094 (.023)	.117 (.025)	-.203 (.022)	-.064 (.025)	.265 (.048)	.409 (.058)
Unweighted	-.100 (.023)	.092 (.024)	-.206 (.021)	-.057 (.024)	.285 (.034)	.387 (.035)
Other test score measures:						
<i>T</i> -scores	-.050 (.024)	.141 (.030)	-.057 (.022)	.065 (.028)	.176 (.040)	.298 (.048)
By gender						
Males	-.126 (.034)	.093 (.037)	-.224 (.032)	-.095 (.035)	.338 (.078)	.385 (.087)
Females	-.058 (.030)	.147 (.035)	-.181 (.031)	-.035 (.036)	.203 (.059)	.433 (.077)
By SES quintile:						
Bottom	-.092 (.044)	-.005 (.041)	-.202 (.044)	-.133 (.045)	.328 (.143)	.043 (.111)
Second	-.088 (.045)	.091 (.049)	-.179 (.046)	-.090 (.047)	.044 (.106)	-.001 (.090)
Third	-.097 (.049)	.068 (.045)	-.242 (.046)	-.106 (.051)	.249 (.121)	.351 (.167)
Fourth	-.082 (.058)	.292 (.077)	-.100 (.056)	.030 (.057)	.207 (.088)	.396 (.115)
Top	-.169 (.080)	.068 (.085)	-.323 (.078)	-.113 (.094)	.404 (.087)	.724 (.102)
By family structure:						
Single mother	-.087 (.043)	.070 (.043)	-.197 (.048)	-.119 (.047)	.086 (.149)	.114 (.144)
Two biological parents	-.127 (.034)	.141 (.042)	-.176 (.029)	-.033 (.033)	.291 (.054)	.456 (.064)
Teen mother at 1st birth	-.101 (.036)	.014 (.033)	-.199 (.036)	-.127 (.038)	.170 (.105)	.251 (.114)
Teen mother at child's birth	-.062 (.046)	-.021 (.043)	-.196 (.045)	-.105 (.052)	.279 (.141)	.281 (.135)
By region:						
Northeast	-.087 (.060)	.129 (.076)	-.159 (.054)	-.030 (.060)	.305 (.124)	.483 (.156)
Midwest	.004 (.053)	.093 (.057)	-.140 (.064)	-.031 (.061)	.337 (.119)	.562 (.133)
South	-.153 (.032)	.051 (.033)	-.217 (.040)	-.119 (.048)	.154 (.104)	.368 (.111)
West	.098 (.077)	.362 (.095)	-.200 (.044)	-.001 (.048)	.269 (.071)	.353 (.088)
By location type:						
Central city	-.110 (.035)	.147 (.040)	-.235 (.033)	-.073 (.037)	.271 (.061)	.439 (.075)
Suburban	-.135 (.039)	.030 (.041)	-.261 (.041)	-.145 (.042)	.146 (.102)	.310 (.119)
Rural	-.184 (.048)	-.032 (.050)	-.253 (.062)	-.124 (.072)	.255 (.130)	.126 (.102)
By school type:						
Public	-.106 (.024)	.098 (.027)	-.214 (.024)	-.081 (.027)	.260 (.051)	.392 (.064)
Private	.022 (.070)	.281 (.074)	-.152 (.058)	.015 (.066)	.296 (.135)	.479 (.137)
School >80% black	.053 (.269)	-.016 (.215)	-.084 (.298)	.057 (.273)	.285 (.382)	.788 (.641)
School >80% white	-.105 (.047)	.059 (.053)	-.186 (.025)	-.061 (.028)	.288 (.054)	.436 (.065)

NOTES: Specifications in this table are variations on those reported in columns 4 and 9 of table 2. Only the race coefficients are reported. The top row simply reproduces the baseline results in columns 4 and 9 of table 2. The remaining rows correspond to different weights, test score measures, or particular subsets of the data. For further details of the baseline specification, see the notes to table 2.