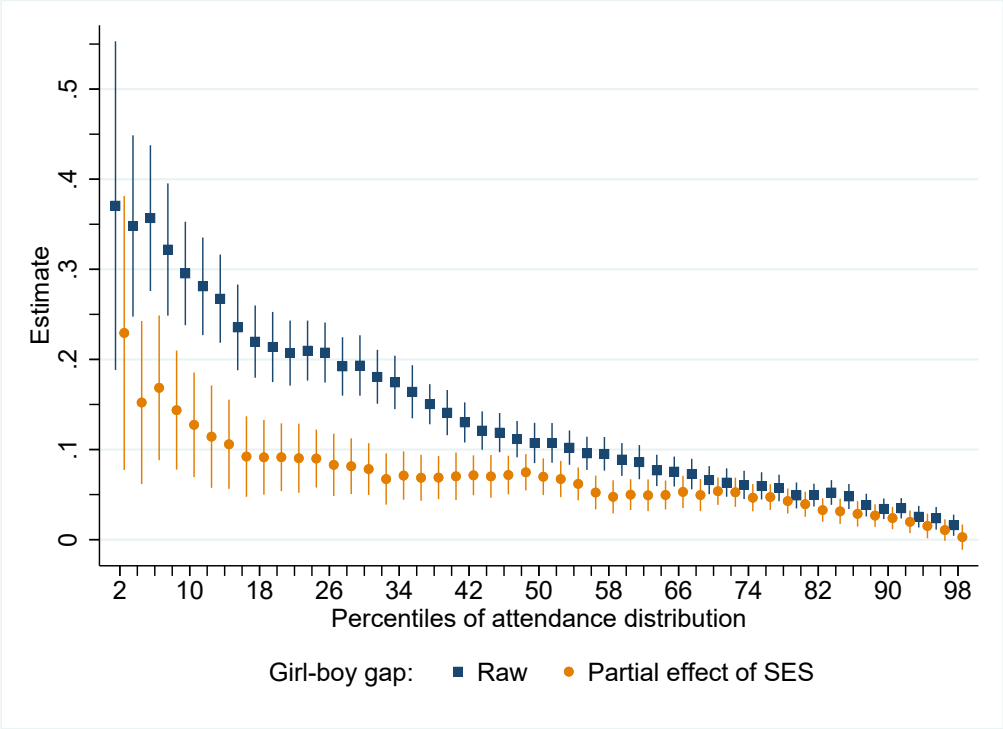
# Appendix for Males at the Tails:

# How Socioeconomic Status Shapes the Gender Gap

David Autor David Figlio Krzysztof Karbownik Jeffrey Roth Melanie Wasserman

# A Supplemental Figures and Tables

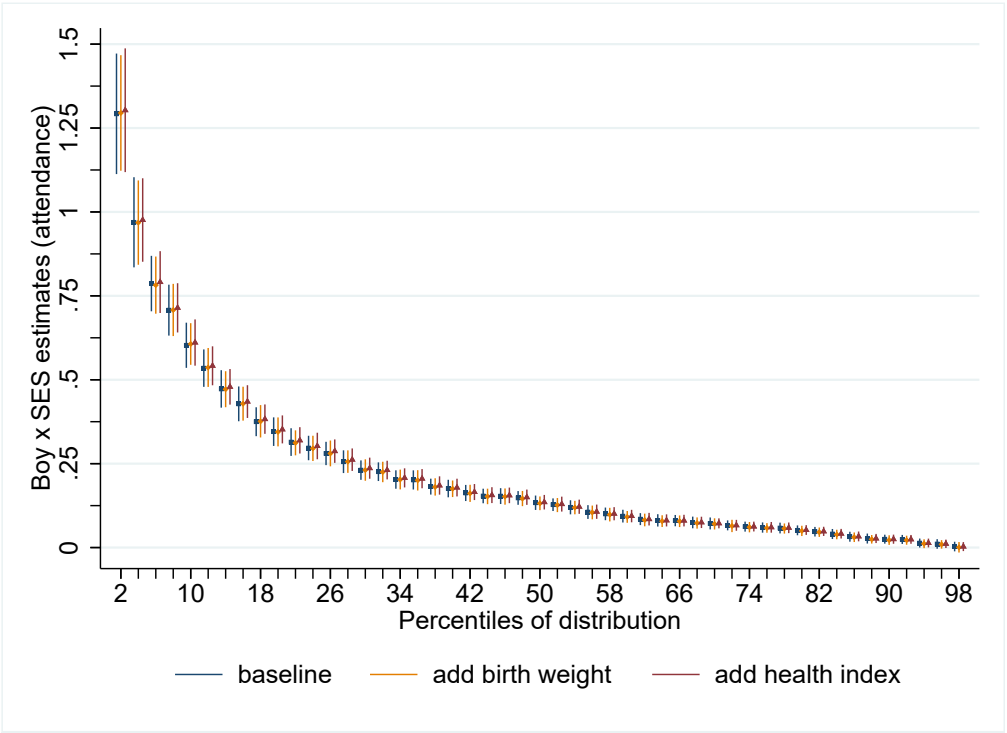
Figure A.1: Effect of SES on the Gender Gap throughout the Distribution: Attendance Variable Based on Absences Only



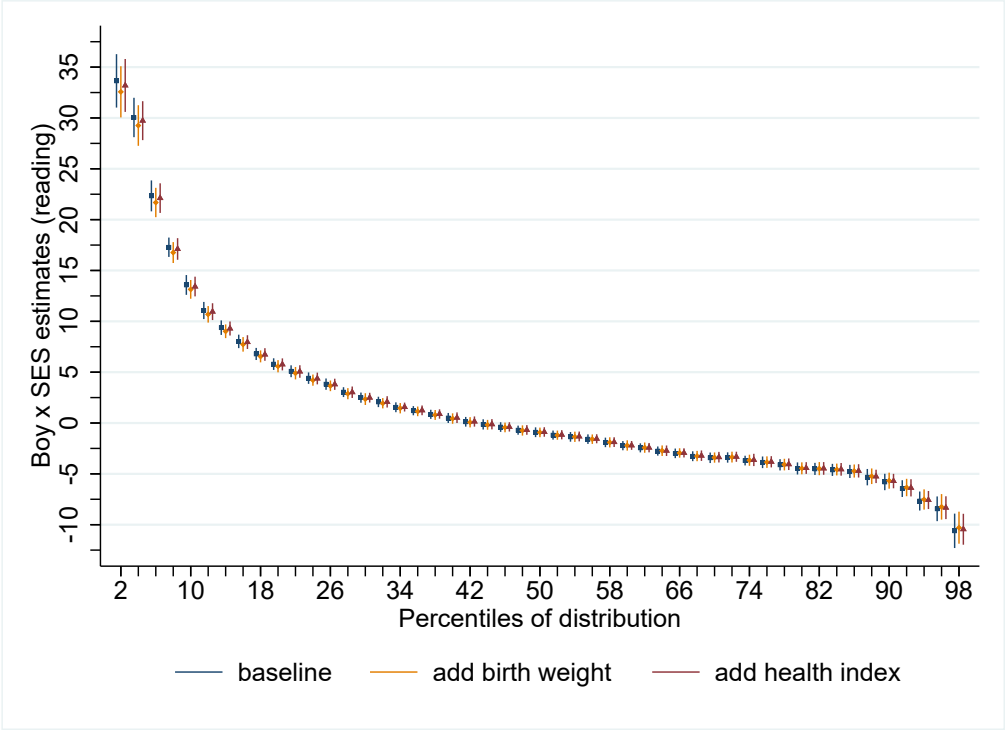

Note: This figure replicates results from Panel A of Figure **??** with a version of attendance that is only based on absences and excludes suspensions. Dependent variable is multiplied by 100. Spikes represent 95% confidence intervals based on bootstrapped standard errors with 200 replications.

Figure A.2: Sensitivity of Main Results to the Inclusion of Neonatal Health

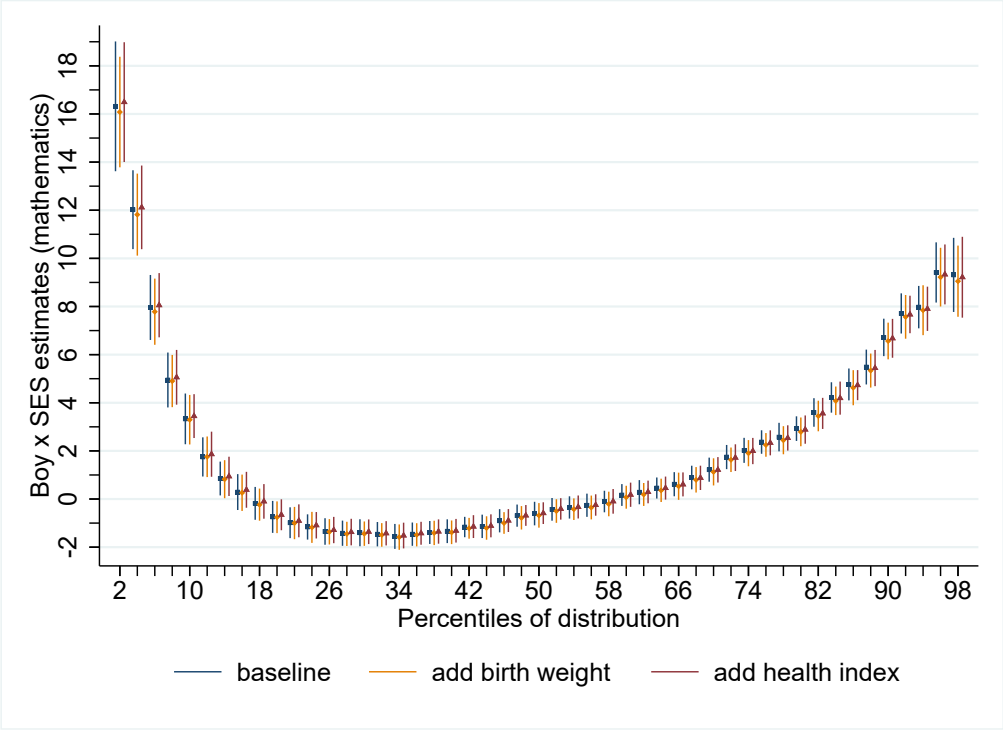
A. Attendance




B. Reading Scores



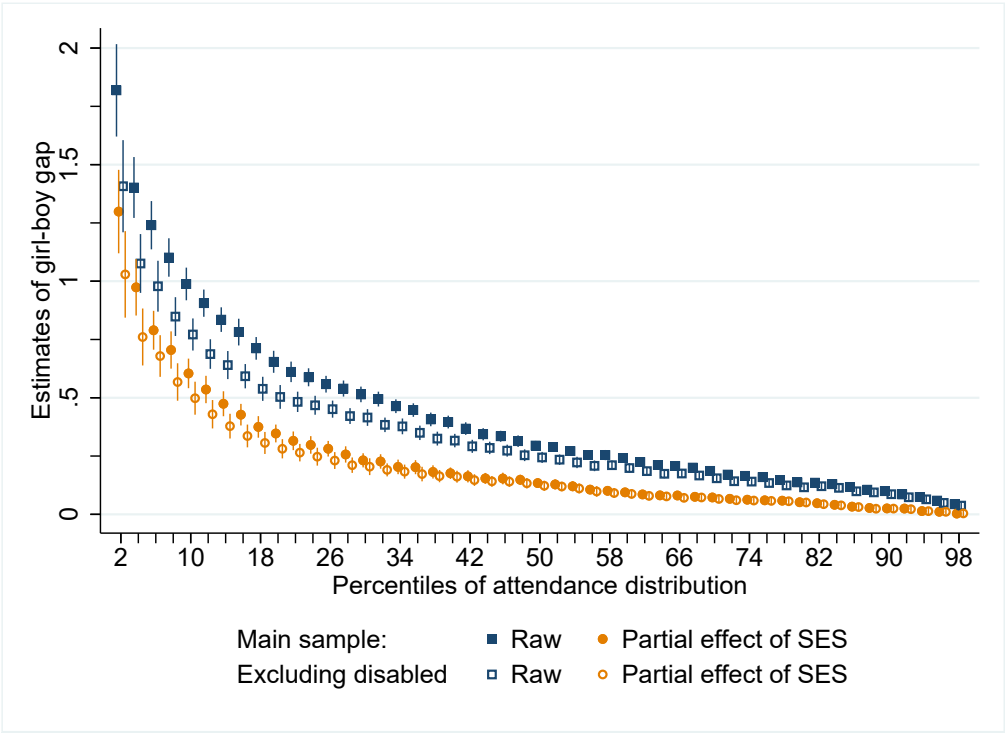
C. Mathematics Scores



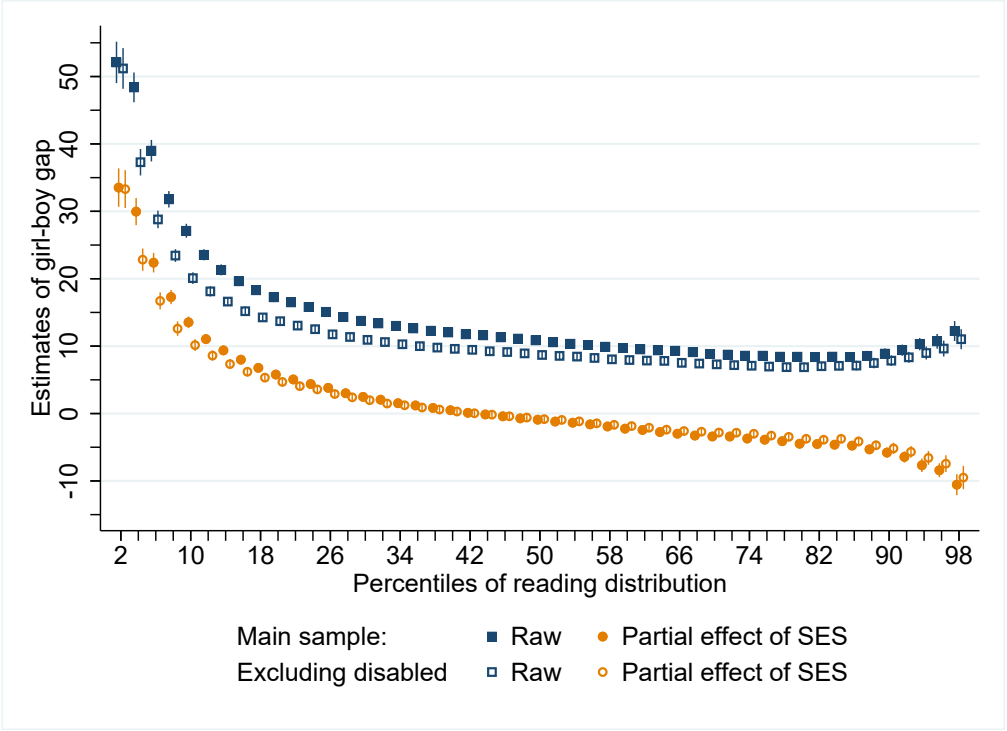

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This figure plots the coefficients on the interaction term Boy x SES, from three separate specifications. The navy squares depict the main results, estimated from equation (**??**). The orange diamonds plots the estimates from a specification that additionally controls for log birth weight and its interaction with Boy. The maroon triangles plot the estimates from a specification that additionally controls for the health index and its interaction with Boy. The health index is based on the first component of a principal components analysis (PCA) using the following variables: birth weight, gestational age, one and five minutes Apgar scores, indicator for adequate prenatal care, indicator for maternal health problems in pregnancy, indicator for congenital disorders, indicator of labour and delivery complications, and indicator for abnormal conditions at birth. Dependent variables are multiplied by 100. Spikes represent 95% confidence intervals based on bootstrapped standard errors with 200 replications.

Figure A.3: Sensitivity of Main Results to the Exclusion of Children with Disabilities

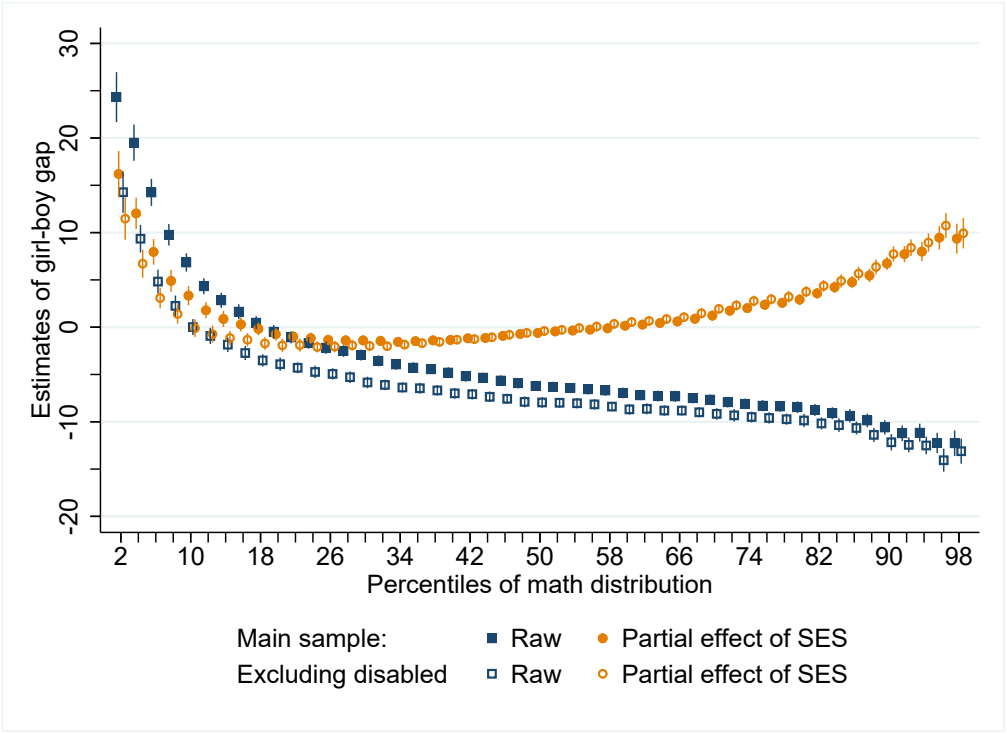
A. Attendance




B. Reading Scores



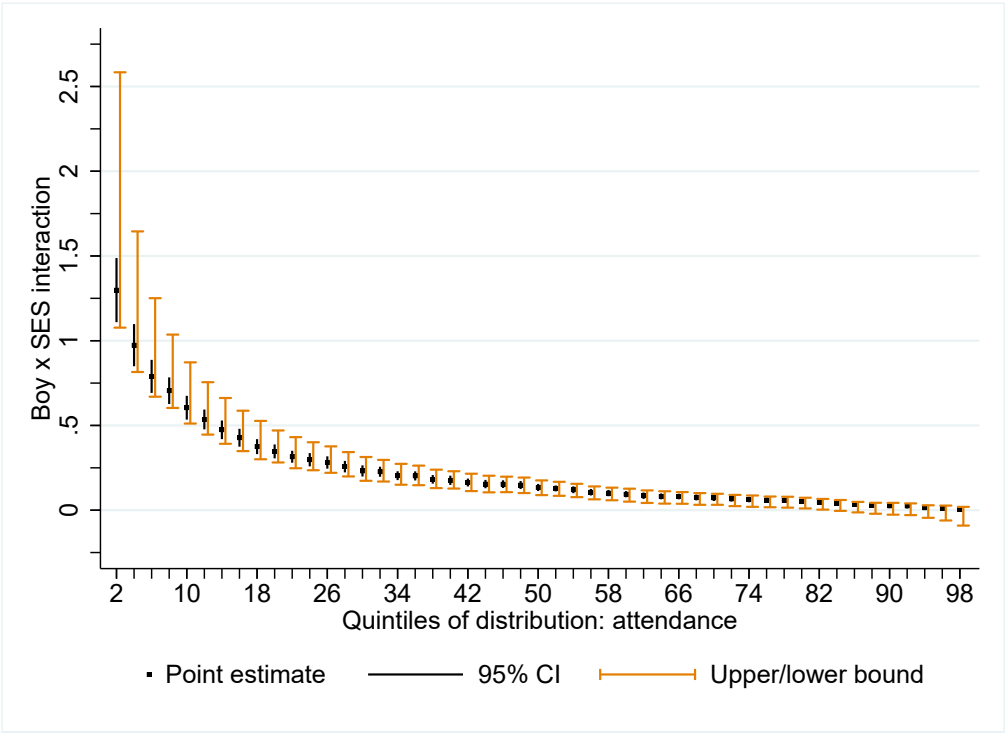
C. Mathematics Scores



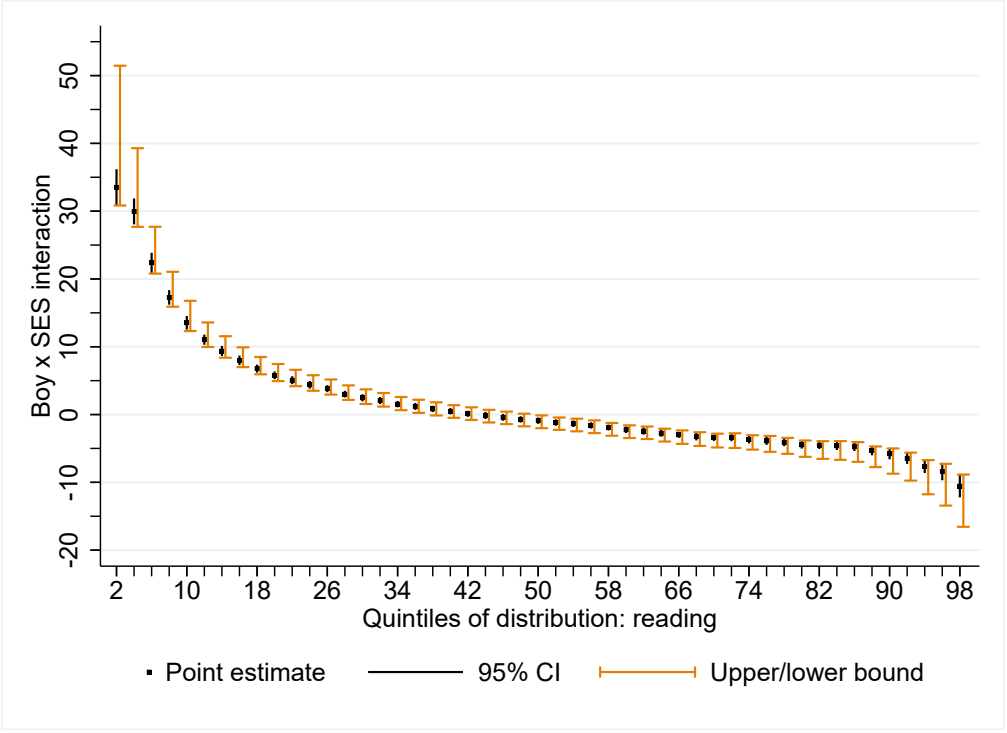

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. The figure plots girl-boy gender gaps (navy squares) and the differential effects of SES on boys (orange circles) for every other percentile of the grades 3 to 8 outcomes' distributions. Dependent variables are multiplied by 100. Panel A presents results for attendance, panel B for reading and panel C for mathematics. The estimates are from RIF regressions (Firpo et al., 2009), estimated using the rifreg command in Stata. Each scatterplot series contains 49 estimates. Raw gender gap estimates come from regressing the outcome on a boy indicator and multiplying these coefficients by -1. Partial effects of SES are the coefficients on Boy x SES interaction terms from a regression of a mathematics test score outcome on boy indicators, SES index, interaction between boy and SES (the plotted coefficients of interest), race/ethnicity indicators, month of birth indicators, year of birth indicators, and birth order indicators. Solid squares and circles replicate our main results from Figure 2 while hollow squares and circles present results from a sample where we exclude 11.3 percent of children with moderate or severe disabilities. Spikes present 95% confidence intervals based on bootstrapped standard errors with 200 replications.

Figure A.4: Bounding the Effects of Gender Imbalance

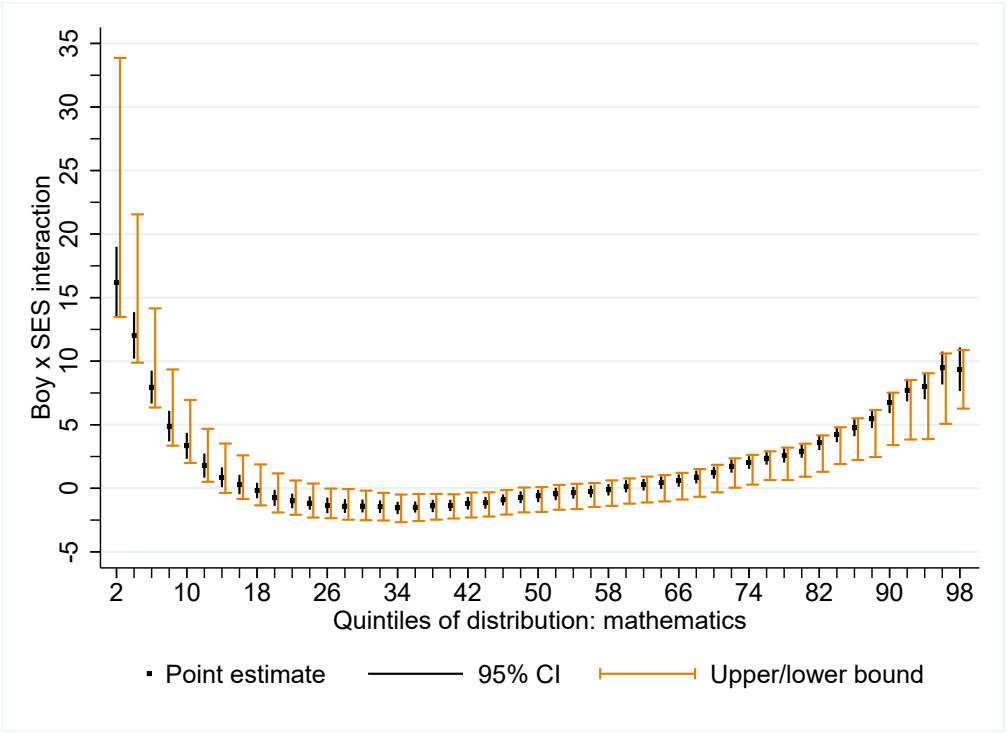
A. Attendance




B. Reading Scores



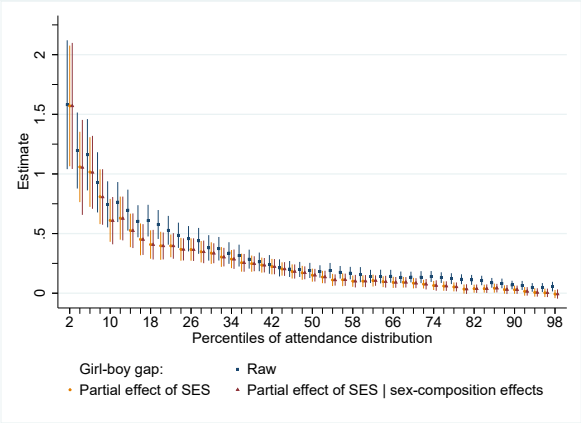
C. Mathematics Scores



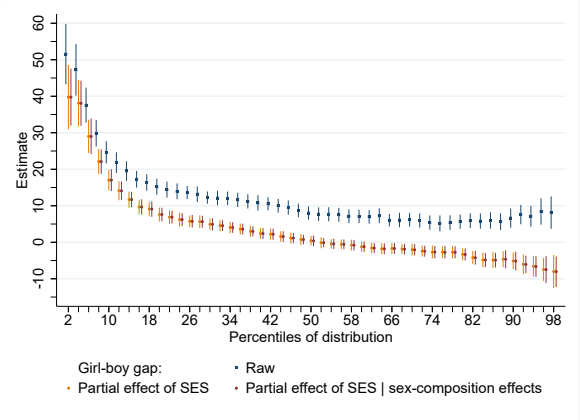

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This figure plots the coefficients on the interaction term Boy x SES, estimated from equation (**??**), and provides Lee bounds for the estimates. Dependent variables are multiplied by 100. See Appendix B for details of the bounding procedure.

Figure A.5: Effect of SES on the Gender Gap throughout the Distribution: First- and second-born siblings sample with controls for sibling-sex composition

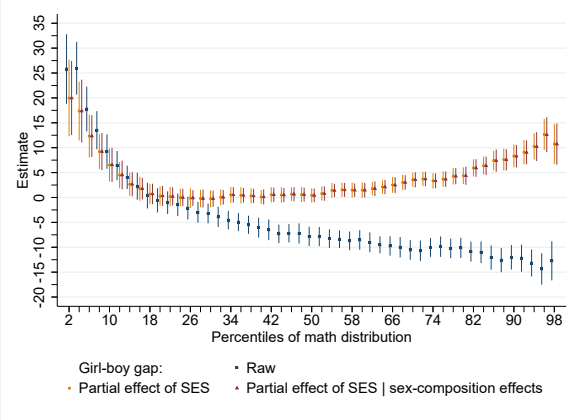
A. Attendance




B. Reading Scores




C. Mathematics Scores




Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes who additionally come from families where we observe at least first- and second-born child. Empirical sample is limited to first-born boys and girls. These figures plot girl-boy gender gaps (navy squares) and the differential effects of SES on boys (orange circles and maroon triangles) for every other percentile of the attendance (Panel A), reading test score (Panel B), and mathematics test score (Panel C) distributions. Dependent variables are multiplied by 100. The estimates are from RIF regressions (**?**) and implemented using the rifreg command in Stata. Each scatterplot series contains 49 estimates. Raw gender gap estimates come from regressing one of the three outcome variables on a Boy indicator and multiplying these coefficients by $-1$. Partial effects of SES (orange circles) are the coefficients on boy x SES interaction terms from a regression of one of the three outcome variables on Boy indicator, SES index, interaction between Boy and SES (the plotted coefficients of interest), race/ethnicity indicators, month of birth indicators, and year of birth indicators. Partial effects of SES conditional on sex-composition effects replicate the previous series but additionally include an indicator for second-born brother and an interaction between indicators for first-born boy and second-born brother as controls. Spikes represent 95% confidence intervals based on bootstrapped standard errors with 200 replications.

Table A.1: Construction of Principal Components SES Index

| | *First component* | *Second component* |
|---|---|---|
| Mother's years of education | 0.49 | 0.71 |
| Married | 0.50 | -0.67 |
| Non-medicaid birth | 0.53 | -0.18 |
| Mother's age at birth | 0.49 | 0.17 |
| Eigenvalue | 2.27 | 0.64 |
| *Summary statistics for the first component* | | |
| Mean | 0.00 | |
| Standard deviation | (1.51) | |
| Mean boys | 0.00 | |
| Standard deviation boys | (1.51) | |
| Mean girls | -0.00 | |
| Standard deviation girls | (1.51) | |

Note: This table reports the results of a principal components analysis of mother's education (in years), mother's age at birth (in years), a non-Medicaid birth indicator, and an indicator for parents married at the time of birth. The eigenvectors associated with the first and second components are reported, as well as their associated eigenvalues. The bottom panel reports summary statistics of the SES index, defined as the first component of the principal components analysis, for the overall sample, and separately for boys and girls.

Table A.2: Sample Selection: Matched Florida Birth and Public School Records

| | White, Black and Hispanic Births | With Complete Data | Matched to Florida School Records | Matched to Outcomes |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| White non-Hispanic | 68.5 | 68.4 | 66.0 | 64.4 |
| Black non-Hispanic | 21.5 | 21.6 | 23.9 | 25.6 |
| Hispanic | 10.0 | 10.0 | 10.1 | 10.0 |
| High school dropout | 18.9 | 18.8 | 20.6 | 21.9 |
| High school graduate | 60.5 | 60.6 | 62.0 | 62.6 |
| College graduate | 20.3 | 20.6 | 17.4 | 15.5 |
| Age 21 or below | 24.1 | 24.0 | 26.1 | 27.4 |
| Age between 22 and 29 | 42.1 | 42.1 | 42.2 | 42.0 |
| Age between 30 and 35 | 24.9 | 25.0 | 23.5 | 22.7 |
| Age 36 or above | 8.9 | 8.9 | 8.3 | 7.9 |
| Married at time of birth | 64.3 | 64.5 | 61.4 | 59.2 |
| Boy | 51.3 | 51.3 | 51.0 | 50.5 |
| Birth weight (grams) | 3,343 | 3,344 | 3,331 | 3,326 |
| N | 940,609 | 900,801 | 734,074 | 552,819 |

Note: This table reports summary statistics (means) for the Florida statewide data for individuals born between 1994 and 2000. The first column is the full sample of Florida births 1994-2000, excluding immigrant mothers; the second column drops the 4.2% of records that are missing key variables; the third column contains the approximately 81% of column 2 records that were matched to Florida school records; and the fourth column is the subset of column 3 for children who remained in Florida public schools through third grade and had at least one outcome. All demographic characteristics are derived from the birth certificate. High school graduates include GED recipients.

Table A.3: Summary Statistics

| | (1) All | (2) Males | (3) Females | (4) Difference (M-F) |
|---|---|---|---|---|
| Attendance (raw) | 94.64 | 94.42 | 94.86 | -0.45 |
| | (4.51) | (4.75) | (4.24) | (0.01) |
| Attendance (adjusted) | -0.32 | -0.54 | -0.08 | -0.46 |
| | (4.52) | (4.75) | (4.25) | (0.01) |
| Mathematics | 6.56 | 8.30 | 4.77 | 3.53 |
| | (89.83) | (93.31) | (86.08) | (0.24) |
| Reading | 8.80 | 1.43 | 16.33 | -14.90 |
| | (88.23) | (91.69) | (83.89) | (0.24) |
| Observations | 552,819 | 279,352 | 273,467 | 552,819 |
| High school graduate | 70.66 | 66.93 | 74.26 | -7.33 |
| | (45.53) | (47.05) | (43.72) | (0.24) |
| High school dropout | 16.49 | 18.03 | 15.00 | 3.03 |
| | (37.11) | (38.44) | (35.71) | (0.19) |
| Observations | 144,945 | 71,140 | 73,805 | 144,945 |

Note: This table reports summary statistics (means, standard deviations, and standard errors) for outcomes of interest in the empirical samples used in the main analyses. Variables in rows 1 to 4 are based on individuals born in Florida between 1994 and 2000 whom we observe with at least one year of outcomes (column 4 of Appendix Table A.2). Variables in rows 5 and 6 are based on individuals born in Florida between 1992 and 1993 for whom we observe high school graduation outcomes. Column 1 presents means and standard deviations for children of both genders; column 2 presents means and standard deviations for males; column 3 presents means and standard deviations for females; while column 4 presents differences in means between columns 2 and 3 with standard errors. Outcome variables are multiplied by 100.

Table A.4: Distributional Effects of Family SES on the Gender Gap: Attendance

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | *A. OLS* | | | *B. Q10* | | | *C. Q25* | | |
| Boy × Family SES | | | 0.25*** | | | 0.60*** | | | 0.29*** |
| | | | (0.01) | | | (0.03) | | | (0.02) |
| Family SES | | 1.32*** | 1.20*** | | 2.49*** | 2.18*** | | 1.84*** | 1.69*** |
| | | (0.01) | (0.01) | | (0.03) | (0.03) | | (0.02) | (0.02) |
| Boy | -0.46*** | -0.49*** | -0.49*** | -0.99*** | -1.05*** | -1.05*** | -0.57*** | -0.62*** | -0.62*** |
| | (0.01) | (0.01) | (0.01) | (0.04) | (0.04) | (0.04) | (0.02) | (0.02) | (0.02) |
| | *D. Q50* | | | *E. Q75* | | | *F. Q90* | | |
| Boy × Family SES | | | 0.13*** | | | 0.06*** | | | 0.03*** |
| | | | (0.01) | | | (0.01) | | | (0.01) |
| Family SES | | 1.13*** | 1.06*** | | 0.65*** | 0.62*** | | 0.38*** | 0.37*** |
| | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.00) | (0.01) |
| Boy | -0.30*** | -0.32*** | -0.32*** | -0.16*** | -0.17*** | -0.17*** | -0.10*** | -0.10*** | -0.10*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Child & mother controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the attendance rate, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 reports the raw gender gap in attendance; columns 2 include the following control variables: family SES index, dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies; columns 3 further include the Boy x SES interaction. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.5: Distributional Effects of Family SES on the Gender Gap: Reading Scores

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | *A. OLS* | | | *B. Q10* | | | *C. Q25* | | |
| Boy × Family SES | | | 1.72*** | | | 13.53*** | | | 4.07*** |
| | | | (0.21) | | | (0.47) | | | (0.24) |
| Family SES | | 29.77*** | 28.90*** | | 27.45*** | 20.59*** | | 26.08*** | 24.02*** |
| | | (0.12) | (0.15) | | (0.29) | (0.33) | | (0.22) | (0.23) |
| Boy | -14.90*** | -16.35*** | -16.36*** | -27.11*** | -28.68*** | -28.70*** | -15.39*** | -16.90*** | -16.90*** |
| | (0.24) | (0.21) | (0.21) | (0.49) | (0.49) | (0.45) | (0.34) | (0.30) | (0.30) |
| | *D. Q50* | | | *E. Q75* | | | *F. Q90* | | |
| Boy × Family SES | | | -0.91*** | | | -3.73*** | | | -5.79*** |
| | | | (0.24) | | | (0.26) | | | (0.36) |
| Family SES | | 30.26*** | 30.72*** | | 31.62*** | 33.51*** | | 32.15*** | 35.08*** |
| | | (0.19) | (0.22) | | (0.23) | (0.24) | | (0.30) | (0.37) |
| Boy | -10.87*** | -12.41*** | -12.41*** | -8.61*** | -10.00*** | -9.99*** | -8.89*** | -10.14*** | -10.13*** |
| | (0.30) | (0.26) | (0.27) | (0.30) | (0.29) | (0.27) | (0.36) | (0.41) | (0.35) |
| Child & mother controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the standardised reading score, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 reports the raw gender gap in attendance; columns 2 include the following control variables: family SES index, dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies; columns 3 further include the Boy x SES interaction. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.6: Distributional Effects of Family SES on the Gender Gap: Math Scores

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | *A. OLS* | | | *B. Q10* | | | *C. Q25* | | |
| Boy × Family SES | | | 2.28*** | | | 3.34*** | | | -1.33*** |
| | | | (0.21) | | | (0.51) | | | (0.33) |
| Family SES | | 30.00*** | 28.85*** | | 28.03*** | 26.34*** | | 27.30*** | 27.98*** |
| | | (0.12) | (0.15) | | (0.35) | (0.39) | | (0.23) | (0.25) |
| Boy | 3.53*** | 2.05*** | 2.05*** | -6.84*** | -8.61*** | -8.61*** | 1.98*** | 0.40 | 0.40 |
| | (0.24) | (0.21) | (0.21) | (0.46) | (0.49) | (0.52) | (0.35) | (0.31) | (0.32) |
| | *D. Q50* | | | *E. Q75* | | | *F. Q90* | | |
| Boy × Family SES | | | -0.60** | | | 2.29*** | | | 6.74*** |
| | | | (0.24) | | | (0.27) | | | (0.41) |
| Family SES | | 29.85*** | 30.15*** | | 32.02*** | 30.86*** | | 33.02*** | 29.61*** |
| | | (0.19) | (0.24) | | (0.22) | (0.24) | | (0.30) | (0.34) |
| Boy | 6.21*** | 4.72*** | 4.72*** | 8.19*** | 6.81*** | 6.81*** | 10.60*** | 9.35*** | 9.34*** |
| | (0.29) | (0.23) | (0.25) | (0.31) | (0.29) | (0.28) | (0.36) | (0.34) | (0.36) |
| Child & mother controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the standardised mathematics score, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 reports the raw gender gap in attendance; columns 2 include the following control variables: family SES index, dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies; columns 3 further include the Boy x SES interaction. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.7: Relationship between Intermediate Outcomes and High School Outcomes

| | Attendance | Attendance squared | Attendance cubed | Math | Math squared | Math cubed | Reading | Reading squared | Reading cubed |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Panel A. High school dropout | | | | | |
| Outcome: High school dropout | -11.98*** | -0.27* | 0.05*** | -5.23*** | 0.22*** | 0.42*** | -1.80*** | 0.32*** | 0.23*** |
| Mean of Y = 16.49; N = 144,945 | (0.16) | (0.14) | (0.02) | (0.22) | (0.08) | (0.03) | (0.22) | (0.08) | (0.04) |
| Marginal effect at 10th percentile | | -11.11 | | | -3.91 | | | -1.53 | |
| Marginal effect at 25th percentile | | -11.77 | | | -5.11 | | | -1.95 | |
| Marginal effect at 50th percentile | | -12.11 | | | -5.18 | | | -1.76 | |
| Marginal effect at 75th percentile | | -12.27 | | | -4.38 | | | -1.09 | |
| Marginal effect at 90th percentile | | -12.34 | | | -3.06 | | | -0.09 | |
| | | | | Panel B: High school graduation | | | | | |
| Outcome: High school graduation | 16.81*** | 1.52*** | 0.02** | 9.52*** | -0.75*** | -0.68*** | 4.00*** | -0.62*** | -0.46*** |
| Mean of Y = 70.66; N = 144,945 | (0.17) | (0.08) | (0.01) | (0.26) | (0.09) | (0.04) | (0.26) | (0.09) | (0.04) |
| Marginal effect at 10th percentile | | 13.22 | | | 8.37 | | | 3.39 | |
| Marginal effect at 25th percentile | | 15.75 | | | 9.75 | | | 4.25 | |
| Marginal effect at 50th percentile | | 17.66 | | | 9.39 | | | 3.92 | |
| Marginal effect at 75th percentile | | 18.86 | | | 7.64 | | | 2.58 | |
| Marginal effect at 90th percentile | | 19.58 | | | 5.13 | | | 0.60 | |

Note: This table reports coefficients and marginal effects from OLS estimation of equation **??**. We regress an indicator for high school dropout (Panel A) or high school graduation (Panel B) on cubic polynomials of standardised attendance, mathematics test scores, and reading test scores. Standardised input variables have mean 0 and SD of 1 in the empirical sample. Each panel reports estimates from a separate regression with 9 displayed variables and additional controls. Outcome variables are multiplied by 100. Each row reports marginal effects for the three intermediate inputs (attendance, mathematics, reading) at the 10th, 25th, 50th, 75th, and 90th percentiles. Additional controls include: SES index, and indicators for gender, race/ethnicity, month and year of birth, birth order, and an interaction between SES and Boy. Heteroskedasticity robust standard errors are in parentheses.

Table A.8: High School Outcomes at Deciles of Intermediate Outcome Distribution

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| | Deciles of intermediate outcomes distribution | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | Panel A. Attendance | | | | | | | | |
| High school dropout | 37.7 | 25.9 | 18.9 | 14.7 | 11.7 | 9.5 | 7.9 | 6.4 | 4.9 |
| High school graduation | 40.1 | 55.1 | 64.9 | 71.7 | 76.2 | 80.3 | 83.3 | 86.2 | 88.9 |
| Boy-girl gender gap in high school dropout | 6.8 | 5.0 | 3.3 | 3.8 | 3.0 | 2.2 | 1.6 | 1.3 | 0.8 |
| Boy-girl gender gap in high school graduation | -12.2 | -10.1 | -8.5 | -8.2 | -8.4 | -7.3 | -5.7 | -3.9 | -3.1 |
| | Panel B. Mathematics | | | | | | | | |
| High school dropout | 29.6 | 24.4 | 21.2 | 18.4 | 15.2 | 12.9 | 11.5 | 8.8 | 6.7 |
| High school graduation | 46.2 | 56.3 | 62.8 | 68.2 | 73.2 | 77.2 | 80.1 | 84.5 | 88.9 |
| Boy-girl gender gap in high school dropout | 6.6 | 5.9 | 4.8 | 4.6 | 3.0 | 2.0 | 1.5 | 1.2 | 0.9 |
| Boy-girl gender gap in high school graduation | -12.4 | -11.3 | -10.9 | -10.5 | -8.0 | -5.9 | -4.9 | -4.0 | -2.5 |
| | Panel C. Reading | | | | | | | | |
| High school dropout | 27.6 | 23.3 | 19.8 | 17.7 | 15.8 | 13.2 | 11.7 | 10.1 | 8.2 |
| High school graduation | 48.1 | 58.0 | 65.1 | 68.9 | 72.7 | 76.7 | 79.4 | 82.9 | 86.6 |
| Boy-girl gender gap in high school dropout | 6.6 | 5.0 | 3.3 | 5.2 | 2.6 | 1.8 | 1.3 | 1.8 | -0.1 |
| Boy-girl gender gap in high school graduation | -12.2 | -11.3 | -8.8 | -9.7 | -8.3 | -5.9 | -4.9 | -4.8 | -2.1 |

Note: This table reports high school outcomes at deciles of the pooled male/female intermediate outcome distributions. The gender gap in high school outcomes refers to the boy-girl difference in high school outcomes at deciles of gender-specific intermediate outcome distributions.

Table A.9: Distributional Effects of Family SES on the Gender Gap in Neonatal Health

| | OLS | | Q10 | | Q25 | | Q50 | | Q75 | | Q90 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (1) | (2) | (1) | (2) | (1) | (2) | (1) | (2) | (1) | (2) |
| | | | | | | *A. Ln(Birth Weight)×100* | | | | | | |
| Boy×Family SES | | -0.05 | | -0.98*** | | -0.57*** | | -0.03 | | 0.55*** | | 0.90*** |
| | | (0.05) | | (0.10) | | (0.06) | | (0.05) | | (0.04) | | (0.06) |
| Family SES | | 1.67*** | | 2.42*** | | 2.07*** | | 1.70*** | | 1.22*** | | 0.87*** |
| | | (0.04) | | (0.09) | | (0.05) | | (0.04) | | (0.03) | | (0.04) |
| Boy | 3.70*** | 3.57*** | 3.29*** | 3.13*** | 3.52*** | 3.39*** | 3.72*** | 3.60*** | 3.87*** | 3.76*** | 3.80*** | 3.72*** |
| | (0.05) | (0.05) | (0.13) | (0.12) | (0.07) | (0.07) | (0.05) | (0.05) | (0.06) | (0.05) | (0.07) | (0.06) |
| Mean of Y | | | | | | 809.27 | | | | | | |
| # children | | | | | | 552,819 | | | | | | |
| | | | | | | *B. Health at Birth Index×100* | | | | | | |
| Boy×Family SES | | -0.68** | | -2.51*** | | -1.44*** | | -0.29 | | 0.26 | | 0.99*** |
| | | (0.26) | | (0.70) | | (0.31) | | (0.20) | | (0.18) | | (0.22) |
| Family SES | | 4.83*** | | 9.30*** | | 6.49*** | | 4.07*** | | 2.58*** | | 1.23*** |
| | | (0.20) | | (0.54) | | (0.24) | | (0.16) | | (0.15) | | (0.16) |
| Boy | 2.97*** | 2.56*** | -3.60*** | -4.20*** | 1.37*** | 0.97*** | 4.07*** | 3.75*** | 5.28*** | 5.01*** | 6.09*** | 5.83*** |
| | (0.27) | (0.26) | (0.74) | (0.72) | (0.31) | (0.35) | (0.21) | (0.20) | (0.20) | (0.19) | (0.22) | (0.23) |
| Mean of Y | | | | | | 0.61 | | | | | | |
| # children | | | | | | 552,025 | | | | | | |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variables are log birthweight (in grams) and a neonatal health index at birth based on a principal component analysis of all birth outcomes, including birthweight in grams, gestational age in weeks, one and five minutes Apgar scores, indicator for adequate prenatal care, indicator for maternal health problems in pregnancy, indicator for congenital disorders, indicator of labour and delivery complications, and indicator for abnormal conditions at birth. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 report the raw gender gap in either log birth weight or health index at birth; columns 2 include the following additional control variables: Boy x SES interaction, family SES index, dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies. Abnormal conditions is a dummy variable equal to one if any of the following conditions is observed: anemia; birth injury; fetal alcohol syndrome; hyaline membrane disease; meconium aspiration syndrome; assisted ventilation <30 minutes; assisted ventilation>30 minutes; seizure; or other specified abnormal conditions. Mother health issues during pregnancy is equal to one if the mother suffered from any of a large set of chronic or pregnancy-related disorders (anemia; cardiac disease; acute or chronic lung disease; diabetes; genital herpes, hydramnios/oligohydramnios; hemoglobinopathy; chronic hypertension; pregnancy associated hypertension; eclampsia; incompetent cervix; previous infant 4000+ grams; previous preterm or small for gestational age infant; renal disease; RH sensitization; uterine bleeding; other specified health problem) during pregnancy or delivery. Congenital anomaly indicator is equal to one if a child has been diagnosed with any of a large set of congenital conditions: anencephalus, spina bifida, hydrocephalus, microcephalus, other central nervous system anomalies, head malformations, other circulatory/respiratory anomalies, rectal atresia, esophageal fistula, gastroschisis, other gastrointestinal anomalies, malformed genitalia, renal agenisis, other urogenital anomalies, cleft lip, dactyly issues, club foot, diaphragmatic hernia, other musculoskeletal anomalies, Downs Syndrome, or other chromosomal anomalies. Complications of labour and delivery indicator is equal to one if the birth suffered from any of the following conditions: fever, heavy meconium, premature rupture of membranes, abruptio placenta, placenta previa, other excessive bleeding, seizures during labour, precipitous labour, prolonged labour, dysfunctional labour, breech, cephalopelvic disproportion, cord prolapse, anesthetic complications, or fetal distress. Prenatal care adequacy is defined according to the Kessner Adequacy of Prenatal Care Utilization index (APCU), which is equal to one if the mother received standard prenatal care services during pregnancy. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.10: Exogeneity of Gender

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Outcome: Boy*100 | | | |
| | Full population | | Empirical sample | |
| SES index | 0.05 | -0.00 | 0.69*** | 0.50*** |
| | (0.03) | (0.04) | (0.07) | (0.08) |
| Mean of Y | 51.27 | | 50.53 | |
| Observations | 939,810 | | 552,819 | |
| Controls | No | Yes | No | Yes |

Note: This table reports estimates from an OLS specification that regresses an indicator variable, multiplied by 100, for whether a child is a boy on family SES. Columns 1 and 2 use the full sample of births in the state of Florida. Columns 3 and 4 use the sample of births matched to public schooling records. Columns 1 and 3 do not include any additional controls. Columns 2 and 4 include controls for birth month, birth year, birth order, and maternal race/ethnicity dummies. Heteroskedasticity robust standard errors are in parentheses.

Table A.11: Determinants of the Gender Gap throughout the Distribution: Attendance

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | | A. OLS | | | B. Q10 | | | C. Q25 | |
| Boy × Family SES | 0.25*** | 0.18*** | | 0.60*** | 0.44*** | | 0.29*** | 0.23*** | |
| | (0.01) | (0.01) | | (0.03) | (0.04) | | (0.02) | (0.02) | |
| Boy × School | | 0.10*** | 0.14*** | | 0.20*** | 0.32*** | | 0.09*** | 0.15*** |
| Quality | | (0.01) | (0.01) | | (0.04) | (0.04) | | (0.02) | (0.02) |
| Boy × | | 0.06*** | 0.12*** | | 0.16*** | 0.30*** | | 0.04* | 0.12*** |
| Neighborhood SES | | (0.01) | (0.01) | | (0.04) | (0.04) | | (0.02) | (0.02) |
| Family SES | 1.04*** | 1.08*** | 1.17*** | 1.86*** | 1.94*** | 2.17*** | 1.47*** | 1.50*** | 1.62*** |
| | (0.01) | (0.01) | (0.01) | (0.03) | (0.03) | (0.03) | (0.02) | (0.02) | (0.02) |
| School Quality | 0.38*** | 0.34*** | 0.31*** | 0.79*** | 0.69*** | 0.63*** | 0.52*** | 0.48*** | 0.45*** |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.01) | (0.01) | (0.02) |
| Neighborhood SES | 0.04*** | 0.01 | -0.02** | 0.09*** | 0.01 | -0.06** | 0.08*** | 0.06*** | 0.02 |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.03) | (0.03) | (0.01) | (0.02) | (0.02) |
| Boy | -0.49*** | -0.49*** | -0.49*** | -1.04*** | -1.04*** | -1.04*** | -0.61*** | -0.61*** | -0.61*** |
| | (0.01) | (0.01) | (0.01) | (0.04) | (0.04) | (0.03) | (0.02) | (0.02) | (0.02) |
| | | D. Q50 | | | E. Q75 | | | F. Q90 | |
| Boy × Family SES | 0.13*** | 0.09*** | | 0.06*** | 0.03*** | | 0.03*** | 0.00 | |
| | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) | |
| Boy × School | | 0.06*** | 0.08*** | | 0.04*** | 0.05*** | | 0.03*** | 0.03*** |
| Quality | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) |
| Boy × | | 0.04*** | 0.07*** | | 0.03*** | 0.04*** | | 0.02** | 0.02** |
| Neighborhood SES | | (0.01) | (0.01) | | (0.01) | (0.01) | | (0.01) | (0.01) |
| Family SES | 0.94*** | 0.96*** | 1.01*** | 0.57*** | 0.58*** | 0.60*** | 0.35*** | 0.36*** | 0.36*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | (0.01) | (0.01) | (0.00) |
| School Quality | 0.30*** | 0.27*** | 0.26*** | 0.16*** | 0.14*** | 0.13*** | 0.08*** | 0.07*** | 0.07*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | (0.01) | (0.01) |
| Neighborhood SES | 0.04*** | 0.02** | 0.00 | 0.00 | -0.01** | -0.02*** | -0.02*** | -0.03*** | -0.03*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | (0.01) | (0.01) |
| Boy | -0.31*** | -0.31*** | -0.31*** | -0.17*** | -0.17*** | -0.17*** | -0.10*** | -0.10*** | -0.10*** |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the attendance rate, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 report results based on navy bars estimates in Figure **??** while columns 2 report results based on orange bars estimates in Figure **??**. In columns 3 we drop the Boy x Family SES interaction terms. All columns additionally include: race/ethnicity indicators, month of birth indicators, year of birth indicators, and birth order indicators. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.12: Determinants of the Gender Gap throughout the Distribution: Reading

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | | A. OLS | | | B. Q10 | | | C. Q25 | |
| Boy × Family SES | 1.71*** | 0.61** | | 13.53*** | 9.15*** | | 4.06*** | 2.51*** | |
| | (0.20) | (0.24) | | (0.51) | (0.54) | | (0.28) | (0.32) | |
| Boy × School Quality | | 1.79*** | 1.96*** | | 5.79*** | 8.23*** | | 2.03*** | 2.70*** |
| | | (0.24) | (0.24) | | (0.45) | (0.54) | | (0.27) | (0.29) |
| Boy × Neighborhood SES | | 0.71*** | 0.91*** | | 4.12*** | 7.05*** | | 1.47*** | 2.28*** |
| | | (0.25) | (0.24) | | (0.53) | (0.58) | | (0.32) | (0.29) |
| Family SES | 23.51*** | 24.07*** | 24.38*** | 15.00*** | 17.23*** | 21.86*** | 18.86*** | 19.65*** | 20.92*** |
| | (0.16) | (0.17) | (0.13) | (0.35) | (0.34) | (0.33) | (0.23) | (0.24) | (0.20) |
| School Quality | 11.82*** | 10.91*** | 10.83*** | 11.96*** | 9.02*** | 7.79*** | 11.61*** | 10.58*** | 10.24*** |
| | (0.12) | (0.17) | (0.16) | (0.26) | (0.34) | (0.32) | (0.16) | (0.21) | (0.23) |
| Neighborhood SES | 3.07*** | 2.71*** | 2.61*** | 3.47*** | 1.36*** | -0.13 | 2.61*** | 1.86*** | 1.45*** |
| | (0.13) | (0.17) | (0.17) | (0.27) | (0.35) | (0.37) | (0.19) | (0.24) | (0.23) |
| Boy | -16.20*** | -16.21*** | -16.21*** | -28.54*** | -28.56*** | -28.56*** | -16.75*** | -16.75*** | -16.75*** |
| | (0.20) | (0.20) | (0.20) | (0.51) | (0.52) | (0.52) | (0.31) | (0.26) | (0.30) |
| | | D. Q50 | | | E. Q75 | | | F. Q90 | |
| Boy × Family SES | -0.92*** | -1.17*** | | -3.73*** | -3.27*** | | -5.80*** | -5.00*** | |
| | (0.23) | (0.29) | | (0.26) | (0.32) | | (0.38) | (0.41) | |
| Boy × School Quality | | 1.04*** | 0.73*** | | 0.16 | -0.72** | | -0.44 | -1.78*** |
| | | (0.28) | (0.27) | | (0.36) | (0.34) | | (0.50) | (0.51) |
| Boy × Neighborhood SES | | -0.43 | -0.80*** | | -1.16*** | -2.21*** | | -1.34*** | -2.94*** |
| | | (0.29) | (0.28) | | (0.33) | (0.36) | | (0.44) | (0.44) |
| Family SES | 25.34*** | 25.47*** | 24.88*** | 28.20*** | 27.96*** | 26.30*** | 29.63*** | 29.22*** | 26.69*** |
| | (0.23) | (0.25) | (0.21) | (0.25) | (0.30) | (0.20) | (0.32) | (0.39) | (0.28) |
| School Quality | 11.78*** | 11.26*** | 11.42*** | 11.68*** | 11.60*** | 12.05*** | 12.13*** | 12.36*** | 13.03*** |
| | (0.15) | (0.22) | (0.22) | (0.19) | (0.27) | (0.25) | (0.23) | (0.36) | (0.36) |
| Neighborhood SES | 3.06*** | 3.27*** | 3.46*** | 2.98*** | 3.57*** | 4.11*** | 2.92*** | 3.59*** | 4.41*** |
| | (0.15) | (0.23) | (0.21) | (0.18) | (0.25) | (0.28) | (0.22) | (0.31) | (0.33) |
| Boy | -12.25*** | -12.25*** | -12.25*** | -9.84*** | -9.84*** | -9.84*** | -9.97*** | -9.97*** | -9.97*** |
| | (0.26) | (0.27) | (0.25) | (0.28) | (0.24) | (0.27) | (0.35) | (0.34) | (0.34) |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the standardised reading score, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 report results based on navy bars estimates in Figure **??** while columns 2 report results based on orange bars estimates in Figure **??**. In columns 3 we drop the Boy x Family SES interaction terms. All columns additionally include: race/ethnicity indicators, month of birth indicators, year of birth indicators, and birth order indicators. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.13: Determinants of the Gender Gap throughout the Distribution: Math

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | | A. OLS | | | B. Q10 | | | C. Q25 | |
| Boy × Family SES | 2.27*** | 0.81*** | | 3.33*** | 1.95*** | | -1.33*** | -2.01*** | |
| | (0.21) | (0.25) | | (0.50) | (0.56) | | (0.30) | (0.35) | |
| Boy × School Quality | | 1.92*** | 2.13*** | | 1.47*** | 1.99*** | | 0.27 | -0.27 |
| | | (0.25) | (0.24) | | (0.56) | (0.57) | | (0.37) | (0.33) |
| Boy × Neighborhood SES | | 1.39*** | 1.65*** | | 1.64*** | 2.27*** | | 1.23*** | 0.58* |
| | | (0.26) | (0.25) | | (0.61) | (0.57) | | (0.36) | (0.35) |
| Family SES | 22.89*** | 23.63*** | 24.04*** | 20.92*** | 21.62*** | 22.61*** | 22.73*** | 23.07*** | 22.05*** |
| | (0.16) | (0.17) | (0.13) | (0.41) | (0.43) | (0.33) | (0.26) | (0.31) | (0.22) |
| School Quality | 13.07*** | 12.10*** | 11.99*** | 12.52*** | 11.78*** | 11.52*** | 12.10*** | 11.96*** | 12.23*** |
| | (0.13) | (0.17) | (0.17) | (0.30) | (0.40) | (0.39) | (0.19) | (0.25) | (0.24) |
| Neighborhood SES | 3.38*** | 2.67*** | 2.54*** | 2.44*** | 1.60*** | 1.28*** | 2.37*** | 1.75*** | 2.08*** |
| | (0.13) | (0.18) | (0.17) | (0.29) | (0.41) | (0.38) | (0.19) | (0.28) | (0.26) |
| Boy | 2.22*** | 2.21*** | 2.21*** | -8.44*** | -8.45*** | -8.45*** | 0.56** | 0.56* | 0.56** |
| | (0.21) | (0.21) | (0.21) | (0.50) | (0.48) | (0.48) | (0.27) | (0.34) | (0.28) |
| | | D. Q50 | | | E. Q75 | | | F. Q90 | |
| Boy × Family SES | -0.61** | -1.43*** | | 2.28*** | 0.89*** | | 6.74*** | 4.37*** | |
| | (0.25) | (0.28) | | (0.26) | (0.33) | | (0.40) | (0.43) | |
| Boy × School Quality | | 0.88*** | 0.50* | | 2.14*** | 2.38*** | | 4.21*** | 5.37*** |
| | | (0.29) | (0.29) | | (0.34) | (0.31) | | (0.46) | (0.43) |
| Boy × Neighborhood SES | | 0.96*** | 0.50** | | 1.02*** | 1.31*** | | 1.20*** | 2.60*** |
| | | (0.30) | (0.26) | | (0.32) | (0.29) | | (0.45) | (0.42) |
| Family SES | 24.46*** | 24.88*** | 24.16*** | 24.43*** | 25.14*** | 25.59*** | 22.48*** | 23.68*** | 25.89*** |
| | (0.23) | (0.22) | (0.20) | (0.23) | (0.27) | (0.21) | (0.34) | (0.35) | (0.28) |
| School Quality | 12.72*** | 12.28*** | 12.47*** | 13.70*** | 12.62*** | 12.50*** | 14.78*** | 12.66*** | 12.07*** |
| | (0.14) | (0.23) | (0.23) | (0.17) | (0.24) | (0.23) | (0.27) | (0.35) | (0.31) |
| Neighborhood SES | 2.97*** | 2.48*** | 2.71*** | 4.05*** | 3.53*** | 3.38*** | 4.92*** | 4.30*** | 3.58*** |
| | (0.15) | (0.22) | (0.21) | (0.18) | (0.23) | (0.23) | (0.22) | (0.30) | (0.28) |
| Boy | 4.89*** | 4.89*** | 4.89*** | 6.99*** | 6.98*** | 6.98*** | 9.54*** | 9.52*** | 9.52*** |
| | (0.23) | (0.27) | (0.24) | (0.27) | (0.29) | (0.26) | (0.34) | (0.36) | (0.37) |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the standardised mathematics score, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Columns 1 report results based on navy bars estimates in Figure **??** while columns 2 report results based on orange bars estimates in Figure **??**. In columns 3 we drop the Boy x Family SES interaction terms. All columns additionally include: race/ethnicity indicators, month of birth indicators, year of birth indicators, and birth order indicators. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.14: Effects of Family SES on the Gender Gap in Parental School Choices

| | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|---|---|---|
| | *School quality* | | | *Dropout rate* | | | *Graduation rate* | | |
| | | | | Panel A: Grade 9 school | | | | | |
| Boy × | | | 0.02 | | | -0.06*** | | | 0.07** |
| Family SES | | | (0.05) | | | (0.02) | | | (0.03) |
| Family SES | | 5.12*** | 5.11*** | | -1.60*** | -1.57*** | | 2.63*** | 2.60*** |
| | | (0.03) | (0.04) | | (0.01) | (0.02) | | (0.02) | (0.02) |
| Boy | -0.45*** | -0.57*** | -0.57*** | 0.18*** | 0.22*** | 0.22*** | -0.34*** | -0.40*** | -0.41*** |
| | (0.05) | (0.05) | (0.05) | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| | | | | Panel B: Grade 6 school | | | | | |
| Boy × | | | 0.08* | | | | | | |
| Family SES | | | (0.04) | | | | | | |
| Family SES | | 5.28*** | 5.24*** | | | | | | |
| | | (0.02) | (0.03) | | | | | | |
| Boy | -0.42*** | -0.51*** | -0.51*** | | | | | | |
| | (0.05) | (0.04) | (0.04) | | | | | | |
| | | | | Panel C: Grade 1 school | | | | | |
| Boy × | | | 0.04 | | | | | | |
| Family SES | | | (0.04) | | | | | | |
| Family SES | | 5.27*** | 5.25*** | | | | | | |
| | | (0.02) | (0.03) | | | | | | |
| Boy | -0.08* | -0.17*** | -0.17*** | | | | | | |
| | (0.04) | (0.04) | (0.04) | | | | | | |

Note: This table reports estimates from OLS regression models where the dependent variables are school quality, average school-level dropout rate, and average school-level graduation rate. Panel A presents results for school a child attends in grade 9, panel B presents results for school a child attends in grade 6, while panel C presents results for school a child attends in grade 1. Dropout and graduation rates are only available for high schools (grade 9). Average school-level dropout and graduation rates are computed from individual level graduation and dropout rates for birth cohorts 1992 and 1993 and are multiplied by 100. This information is then merged at grade 9 school attended by a specific child. School quality measure is based on state school quality reports and ranges from 0 to 100. For each school we compute school-level gain scores calculated by the Florida Department of Education reflecting schools' average contribution to student outcomes. See **?** for detailed description of this measure. Heteroskedasticity robust standard errors are in parentheses.

Table A.15: Distributional Effects of Components of Family SES on the Gender Gap: Attendance

| | Attendance rate | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *OLS* | *Q10* | *Q25* | *Q50* | *Q75* | *Q90* |
| Boy × Mother education | 0.05*** | 0.12*** | 0.04*** | 0.01* | -0.00 | -0.01* |
| | (0.01) | (0.02) | (0.01) | (0.01) | (0.00) | (0.00) |
| Boy × Married | 0.23*** | 0.49*** | 0.23*** | 0.16*** | 0.09*** | 0.06*** |
| | (0.03) | (0.10) | (0.05) | (0.03) | (0.02) | (0.01) |
| Boy × Non-medicaid birth | 0.36*** | 0.96*** | 0.47*** | 0.19*** | 0.08*** | 0.04*** |
| | (0.03) | (0.08) | (0.05) | (0.03) | (0.02) | (0.02) |
| Boy × Mother's age at birth | -0.01*** | -0.03*** | -0.01*** | -0.00** | 0.00 | -0.00 |
| | (0.00) | (0.01) | (0.00) | (0.00) | (0.00) | (0.00) |
| Mother education | 0.33*** | 0.61*** | 0.48*** | 0.31*** | 0.17*** | 0.09*** |
| | (0.00) | (0.01) | (0.01) | (0.01) | (0.00) | (0.00) |
| Married | 0.37*** | 0.79*** | 0.58*** | 0.28*** | 0.14*** | 0.09*** |
| | (0.02) | (0.06) | (0.04) | (0.02) | (0.01) | (0.01) |
| Non-medicaid birth | 1.09*** | 2.13*** | 1.61*** | 0.94*** | 0.51*** | 0.26*** |
| | (0.02) | (0.06) | (0.04) | (0.02) | (0.01) | (0.01) |
| Mother's age at birth | 0.01*** | -0.01** | -0.00 | 0.01*** | 0.01*** | 0.02*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Boy | -1.08*** | -2.44*** | -1.20*** | -0.52*** | -0.23*** | -0.06 |
| | (0.08) | (0.23) | (0.12) | (0.08) | (0.05) | (0.04) |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the attendance rate, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Each regression is akin to columns 3 in Table A.4 but where we use separate components of the SES index: mother's years of education at birth, mother's marital status at birth, an indicator whether the birth was paid by Medicaid and mother's age at birth. All regressions additionally include following controls: dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.16: Distributional Effects of Components of Family SES on the Gender Gap: Reading Scores

| | Reading score | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *OLS* | *Q10* | *Q25* | *Q50* | *Q75* | *Q90* |
| Boy × Mother education | 0.50*** | 3.56*** | 1.08*** | -0.09 | -0.81*** | -1.71*** |
| | (0.12) | (0.26) | (0.14) | (0.15) | (0.16) | (0.23) |
| Boy × Married | 1.89*** | 9.54*** | 3.24*** | -0.34 | -1.91*** | -1.79** |
| | (0.52) | (1.23) | (0.73) | (0.63) | (0.65) | (0.77) |
| Boy × Non-medicaid birth | 2.85*** | 14.22*** | 4.59*** | -0.43 | -2.54*** | -4.09*** |
| | (0.53) | (1.18) | (0.78) | (0.63) | (0.75) | (0.87) |
| Boy × Mother's age at birth | -0.21*** | -0.33*** | -0.16*** | -0.10** | -0.15*** | -0.16** |
| | (0.04) | (0.10) | (0.06) | (0.05) | (0.05) | (0.08) |
| Mother education | 9.02*** | 6.31*** | 6.94*** | 9.12*** | 10.52*** | 11.89*** |
| | (0.08) | (0.18) | (0.12) | (0.11) | (0.13) | (0.20) |
| Married | 6.98*** | 4.62*** | 6.86*** | 8.18*** | 8.28*** | 7.33*** |
| | (0.36) | (0.80) | (0.53) | (0.45) | (0.50) | (0.61) |
| Non-medicaid birth | 16.58*** | 14.61*** | 16.37*** | 18.92*** | 17.49*** | 15.52*** |
| | (0.36) | (0.76) | (0.48) | (0.46) | (0.55) | (0.63) |
| Mother's age at birth | 0.83*** | 0.37*** | 0.58*** | 0.88*** | 1.09*** | 1.18*** |
| | (0.03) | (0.06) | (0.05) | (0.04) | (0.04) | (0.06) |
| Boy | -19.86*** | -77.83*** | -30.53*** | -8.21*** | 6.50*** | 18.71*** |
| | (1.43) | (3.11) | (1.77) | (1.69) | (1.80) | (2.53) |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the reading score, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Each regression is akin to columns 3 in Table A.5 but where we use separate components of the SES index: mother's years of education at birth, mother's marital status at birth, an indicator whether the birth was paid by Medicaid and mother's age at birth. All regressions additionally include following controls: dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

Table A.17: Distributional Effects of Components of Family SES on the Gender Gap: Math Scores

| | Mathematics scores | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *OLS* | *Q10* | *Q25* | *Q50* | *Q75* | *Q90* |
| Boy × Mother education | 0.32** | 0.24 | -0.79*** | -0.68*** | 0.46** | 1.86*** |
| | (0.13) | (0.28) | (0.17) | (0.16) | (0.18) | (0.25) |
| Boy × Married | 2.11*** | 1.84 | 0.62 | 1.68** | 1.95*** | 3.37*** |
| | (0.54) | (1.35) | (0.79) | (0.69) | (0.63) | (0.80) |
| Boy × Non-medicaid birth | 2.90*** | 5.97*** | 0.36 | 0.52 | 2.25*** | 4.98*** |
| | (0.54) | (1.25) | (0.93) | (0.66) | (0.73) | (0.83) |
| Boy × Mother's age at birth | -0.05 | -0.04 | -0.10* | -0.08* | -0.03 | 0.09 |
| | (0.04) | (0.10) | (0.06) | (0.05) | (0.05) | (0.07) |
| Mother education | 9.76*** | 8.95*** | 8.91*** | 9.75*** | 10.59*** | 10.96*** |
| | (0.09) | (0.20) | (0.13) | (0.12) | (0.14) | (0.18) |
| Married | 6.95*** | 8.77*** | 8.84*** | 7.77*** | 6.26*** | 4.24*** |
| | (0.38) | (0.93) | (0.59) | (0.50) | (0.49) | (0.54) |
| Non-medicaid birth | 19.36*** | 22.02*** | 21.63*** | 20.87*** | 17.93*** | 13.40*** |
| | (0.38) | (0.89) | (0.69) | (0.46) | (0.49) | (0.54) |
| Mother's age at birth | 0.30*** | -0.31*** | 0.10** | 0.41*** | 0.60*** | 0.74*** |
| | (0.03) | (0.07) | (0.05) | (0.04) | (0.04) | (0.05) |
| Boy | -3.41** | -14.75*** | 12.40*** | 14.25*** | -0.55 | -21.06*** |
| | (1.46) | (3.54) | (2.10) | (1.65) | (1.83) | (2.61) |

Note: The sample is individuals born between 1994 and 2000 in Florida, attending Florida public schools, and having non-missing attendance and test score outcomes. This table reports estimates from OLS and unconditional quantile regression models, where the dependent variable is the mathematics score, from grades 3 through 8, multiplied by 100. Unconditional quantile regression estimates for 10, 25, 50, 75 and 90 percentiles are obtained using the rifreg command in Stata, developed based on **?**. Each regression is akin to columns 3 in Table A.6 but where we use separate components of the SES index: mother's years of education at birth, mother's marital status at birth, an indicator whether the birth was paid by Medicaid and mother's age at birth. All regressions additionally include following controls: dummies for maternal race and ethnicity, child's month and year of birth dummies, and birth order dummies. For OLS, heteroskedasticity robust standard errors are in parentheses. For unconditional quantile regression, bootstrapped standard errors with 200 replications are in parentheses.

# B Details of Bounding Procedure

Here we provide details regarding the procedure to bound the effects of differential sex selection into the analytical sample. In the full sample of births and in our analytical sample, we divide each sample into quintiles of the SES distribution. For each of those quintiles we compute the difference in fraction of boys across the two samples. These five values are (starting from the first quintile) 1.4 percentage points, 1.3 percentage points, 0.8 percentage points, 0.3 percentage points, and -0.2 percentage points, implying excess boys in the bottom four quintiles and excess girls in the top quintile. Using our analytical sample we compute outcomes of girls at given percentiles for the bottom four quintiles and for boys in the top quintile. For example, since in the bottom quintile we observe 1.4 percentage points excess of girls we compute their outcomes at 1.4th and 98.6th percentile. We next define upper bound samples as females from the bottom four quintiles and males from the top quintile with outcomes greater than the upper percentile values defined in the prior step. Conversely, we define lower bound samples as females from the bottom four quintiles and males from the top quintile with outcomes less than the lower percentile values defined in the prior step. Finally, we run two regressions: (1) equation (**??**) using a sample excluding the upper bound as defined above; (2) equation (**??**) using a sample excluding the lower bound as defined above. The bounds in Figure A.4 report 95 percent confidence intervals for the coefficient ($\text{Boy}_i \times \text{SES}_i$) from these regressions.