American Economic Journal: Applied Economics 2019, 11(3): 338–381
https://doi.org/10.1257/app.20170571

Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes†

By David Autor, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman*

Boys born to disadvantaged families have higher rates of disciplinary problems, lower achievement scores, and fewer high school completions than girls from comparable backgrounds. Using birth certificates matched to schooling records for Florida children born 1992–2002, we find that family disadvantage disproportionately impedes the pre-market development of boys. The differential effect of family disadvantage on boys is robust to specifications within schools and neighborhoods as well as across siblings within families. Evidence supports that this is the effect of the postnatal environment; family disadvantage is unrelated to the gender gap in neonatal health. We conclude that the gender gap among black children is larger than among white children in substantial part because black children are raised in more disadvantaged families. (*JEL* D91, I24, I32, J13, J15, J16)

The last four decades have witnessed a swift and substantial reversal of the gender gap in educational attainment in the United States and much of the developed world. Between 1970 and 2010, the high school graduation rate of US women rose by 6 percentage points, from 81 to 87 percentage points, while the US male

*Autor: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, E52-438, Cambridge, MA 02142, and National Bureau of Economic Research (email: dautor@mit.edu); Figlio: School of Education and Social Policy, Northwestern University, 2120 Campus Drive, Evanston IL 60208, and National Bureau of Economic Research (email: figlio@northwestern.edu); Karbownik: Department of Economics, Emory University, 1602 Fishburne Drive, Atlanta, GA 30322, and National Bureau of Economic Research (email: krzysztof.karbownik@emory.edu); Roth: University of Florida, 1701 SW 16th Avenue, Gainesville, FL 32608 (email: jeffroth@ufl.edu); Wasserman: Anderson School of Management, UCLA, 110 Westwood Plaza C521, Los Angeles, CA 90095 (email: melanie.wasserman@anderson.ucla.edu). Alexandre Mas was coeditor for this article. We thank Josh Angrist, Marianne Bertrand, Raj Chetty, John Ham, Nathan Hendren, Louis Kaplow, Mikael Lindahl, Jeremy Majerovitz, Richard Murnane, Jessica Pan, Kjell Salvanes, Till von Wachter, two anonymous referees, and numerous seminar participants for valuable suggestions that helped to improve the paper. Autor acknowledges support from the Russell Sage Foundation (grant 85-12-07). Figlio and Roth acknowledge support from the National Science Foundation and the Institute for Education Sciences (CALDER grant), and Figlio acknowledges support from the National Institute of Child Health and Human Development and the Bill and Melinda Gates Foundation. Wasserman acknowledges support from the NSF Graduation Research Fellowship, NIA grant T32-AG000186, and NICHD grant HD007339-30. We are grateful to the Florida Departments of Education and Health for providing the de-identified, matched data used in this analysis. The conclusions expressed in this paper are those of the authors and do not represent the positions of the Florida Departments of Education and Health or those of our funders.

†Go to https://doi.org/10.1257/app.20170571 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

high school graduation rate was unchanged (Murnane 2013).[1] Contemporaneously, women have overtaken and surpassed men in higher education: in 2011, the ratio of female to male college attainment among adults ages 25–34 exceeded unity in 28 of 34 OECD countries, with a median above 1.4 (OECD 2013). Amidst this widely remarked rise in female educational attainment hides a striking and comparatively unremarked puzzle: the female advantage in high school graduation and college attainment is larger and has risen by substantially more among children of minority families. For example, while the overall female advantage in high school completions among US adults ages 20 through 24 was 6.2 percentage points in 2010, it was 4.5 percentage points among whites, 12.2 percentage points among blacks, and 7.8 percentage points among US-born Hispanics (Murnane 2013, table 3).[2] Contemporaneous race gaps in college attainment among young US adults are equally pronounced (Snyder and Dillow 2013, table 104.20).[3]

What accounts for the systematically larger gender gaps in educational outcomes among minorities? One possible explanation—the focus of this paper—is gender differences in the effects of family disadvantage. Specifically, we hypothesize that family disadvantage—meaning low availability of household resources, child-rearing inputs (e.g., nutrition, safety in the home, stimuli), and parental attention—differentially inhibits the behavioral and academic development of boys relative to girls, either because these developmental outcomes are more elastic to family circumstances among boys than girls, or because differential parental investment in girls relative to boys varies inversely with household socioeconomic status (SES). Our goals in this paper are: (i) to test whether family disadvantage exerts a disproportionate negative effect on the educational and behavioral outcomes of school-age boys relative to girls; (ii) to differentiate this hypothesis both from a "fetal origins" alternative as well as from a neighborhood-and-school-quality explanation; and (iii) to utilize the resulting estimates to quantify the degree to which higher rates of family disadvantage among minority populations can partly explain the larger gender gaps in educational outcomes we observe among minorities.

In quantifying the contribution of family disadvantage to the gender gap in behavioral and academic outcomes, we face two principal obstacles: suitable data and a credible empirical strategy. To address the data and measurement challenge, we draw upon a matched database of birth certificate, and academic, disciplinary, and high school graduation records for over one million children born in Florida between 1992 and 2002. Florida is particularly well-suited to this research because it has a large, demographically diverse, and socioeconomically heterogeneous population.

[1] High school graduation rates refer to the status completion rate of US born adults ages 20–24, and they include both traditional high school graduates and GED holders. Thus, 1970 graduation rates refer to cohorts born 1946–1950, and 2010 graduation rates refer to cohorts born 1986–1990.

[2] By comparison, the gap in 1970 was 0 overall, −0.4 percentage points among whites, 5.1 percentage points among blacks, and −2.5 percentage points among US born Hispanics. Thus, the increase in the gap among whites, blacks, and Hispanics in this 40-year period was 4.9, 7.1, and 10.3 percentage points, respectively.

[3] Whereas white women ages 25 to 29 were 22 percent more likely to hold a BA than white males in 2010, the corresponding gap was 55 percent among both blacks and Hispanics. Goldin, Katz, and Kuziemko (2006) document that among the high school graduating class of 1992, the female advantage in BA attainment was far higher among children of families in the bottom two SES quartiles than among the top two quartiles, and, moreover, the gender gap in the lower two quartiles had risen by substantially more than among the upper quartiles in the prior 20 years.

Our longitudinal data offer remarkable detail on family characteristics, infant and maternal health at birth, early educational outcomes (including assessments of kindergarten readiness at the start of formal schooling), third through eighth grade test scores, absenteeism, disciplinary outcomes, and high school graduation for the oldest cohorts in our sample. Since family disadvantage is imperfectly observable even in this rich database, we combine multiple variables to proxy for this underlying construct, focusing particularly on maternal age and education, household poverty, and marital status at the time of birth.

The second obstacle to our inquiry is that family environment is intrinsically confounded with congenital, hereditary, and other environmental factors that likely affect children's outcomes independent of the impact of family environment. For example, highly educated parents reside in safer neighborhoods, enroll their children in higher quality schools, and may have children with above average latent ability. Our challenge is thus to separate the direct impact of family environment (e.g., quality and quantity of parental inputs) from the hereditary and environmental confounds that would likely lead to disparate outcomes among children absent any causal effect of family environment on children's development.

Our empirical approach contrasts the outcomes of boys and girls across more versus less advantaged family settings. This strategy provides valid identification of the differential effect of child-rearing environment on boys relative to girls under two conditions. The first condition is that the gender gap in potential outcomes among boys and girls is uncorrelated with our measures of family environment at the time of birth; that is, any intrinsic genetic or biological advantage that girls may possess at birth relative to boys is not systematically larger or smaller in less relative to more disadvantaged families. Using detailed measures of infant and maternal health, obtained from vital statistics, we offer evidence of the plausibility of this assumption. The second condition is that boys and girls are (on average) exposed to the same family environment. This condition could be violated if, for example, family environment is endogenous to the gender of the child, as suggested by Dahl and Moretti (2008). To limit the extent to which unobservable correlates of family environment are driving the socioeconomic gradient in the gender gap, we provide additional analyses in which we contrast the outcomes of opposite-sex siblings within the same family and find results that are highly consistent with the full sample analysis.

We demonstrate that postnatal factors relating to family disadvantage, above and beyond the neighborhoods children inhabit and the schools they attend, substantially and differentially influence the likelihood that boys thrive relative to girls. We begin by documenting that the cross-race differences in the gender gap in long-term educational attainment emerge early in students' academic trajectories and are apparent in both cognitive and behavioral outcomes. For example, 12.8 percent of Florida public school children are suspended at least once between third and eighth grade. But suspensions are 8.1 percentage points higher among white boys than girls, and an additional 4.5 percentage points higher among black boys relative to black girls. We then implement our primary analysis, which demonstrates that boys born to low-SES families perform worse on standardized tests throughout elementary and middle school, have higher rates of absences and behavioral problems, and are less likely to graduate high school than are girls.[4] These differentials are economically large and explain a substantial share of cross-race group differences in the gender gap. Accounting for the differential effect of family disadvantage on boys relative to girls, reduces the cross-race suspension gap to 1.5 percentage points—meaning that 3.0 percentage points of the observed black-white gender gap is proximately explained by higher levels of disadvantage to which both black boys and girls are exposed. Among families of comparable SES, we would predict the black-white gender gap to be 67 percent smaller. Carrying this exercise forward to longer term outcomes, our estimates of the effect of family disadvantage on the gender gap can explain about one-quarter of the excess black-white shortfall in high school on-time completion among black males versus black females relative to white males versus white females.

While the SES gradient in the gender gap could stem entirely from family environment per se, it likely also reflects the differential effect of neighborhood and school quality—both of which are correlated with family income, education, and marital status—on boys relative to girls. We empirically differentiate among these alternatives by augmenting our main analysis with controls for the non-family environment, including: measures of school quality produced by the Florida Department of Education; estimates of the causal effect of counties of residence on economic mobility produced by Chetty and Hendren (2018b); and school and neighborhood indicators. In all cases, we allow these environmental quality measures to differentially affect outcomes of boys relative to girls. After accounting for the direct effect of family characteristics, we find that neighborhood quality makes a modest additional contribution to the boy-girl outcome gap. School quality matters more, however: low quality schools, as measured by Florida state reports of student test score gains, are particularly disadvantageous for boys. Nevertheless, accounting non-parametrically for the differential impact of schools and neighborhoods on boys relative to girls reduces the estimated impact of family disadvantage on the gender gap in behavioral outcomes by at most one quarter.[5]

An implicit assumption of our analytic approach is that the effect of family disadvantage on the gender gap is comparable across race and ethnic groups—for example, that marital status is equally consequential for the gender gap in downstream outcomes among white, black, and Hispanic children. While we do find some important deviations from this benchmark—in particular, greater family SES has a somewhat smaller moderating effect on the gender gap in suspensions among

[4] While the children observed in our sample are not yet old enough for us to observe adult outcomes such as labor force participation, criminal activity, or high school completion (except for the oldest cohorts), Duckworth and Seligman (2005) and Segal (2013) show that observed differences in disciplinary behaviors later manifest in differences in academic and labor market performance. For the oldest two cohorts in our sample, we also document that adverse behavioral outcomes in middle school are strong predictors of high school non-completion, while, by contrast, early academic outcomes have far less explanatory power.

[5] The sociology literature documents that neighborhoods vary within a zip code, which is the finest geography we observe in our data (see Ingoldsby and Shaw 2002; Ramirez-Valles, Zimmerman, and Juarez 2002; Beyers et al. 2003; Leventhal and Brooks-Gunn 2003; and Karriker-Jaffe et al. 2009, for example). If disadvantaged families occupy differentially low-quality "micro-neighborhoods" within larger neighborhoods, and these micro-neighborhoods differentially affect boys more than girls, this would be an unmeasured confounding pathway through which family disadvantage affects the gender gap in developmental outcomes. To the extent to which this is true, our estimated effects of family disadvantage may overstate the causal effect since family disadvantage would remain confounded with unmeasured neighborhood quality, even conditional on zip code fixed effects.

non-white than among whites, perhaps because schools are more punitive toward minority than nonminority boys—accounting for this non-parallelism does not substantively change our findings, as we discuss and interpret below.

This paper contributes to an active literature examining the emerging gender gap in educational and behavioral outcomes. Buchmann and DiPrete (2006), Goldin, Katz, and Kuziemko (2006), DiPrete and Jennings (2012), Autor and Wasserman (2013), Bertrand and Pan (2013), DiPrete and Buchmann (2013), Fortin, Oreopoulos, and Phipps (2015), Chetty et al. (2016b), and Lundberg (2017) explore the evolution of gender differences in behavioral and educational outcomes in the United States and internationally. Several papers explore specific hypotheses for this gender gap. Becker, Hubbard, and Murphy (2010) theorize that gender differences in the psychic costs of education—primarily differences in the distributions of non-cognitive skills—explain the overtaking of men by women in higher education. Papers by Buchmann and DiPrete (2006), DiPrete and Jennings (2012), Baker and Milligan (2013), Bertrand and Pan (2013), Lundberg (2017), Owens (2016), Prevoo and ter Weel (2014), Riphahn and Schwientek (2015), Woessmann (2015), and Aucejo and James (2017) focus on the role of family and schooling environment in generating observed gender gaps in behavioral and educational outcomes.[6]

Bertrand and Pan (2013) offer the most complete extant analysis of the relationship between family background and gender differences in early childhood outcomes. They document that boys raised in single-parent families exhibit twice the rate of behavioral and disciplinary issues as boys raised in two-parent families and are more than twice as likely to be suspended from school by the eighth grade. Considering home environment, school environment, and parent-child interactions, Bertrand and Pan (2013) find only modest evidence that parental and schooling inputs differ systematically between boys and girls across family types. Their analysis suggests instead that the quality of parental and child-rearing inputs has larger impacts on the behavioral and disciplinary outcomes of boys than it does of girls. Consistent with this line of argument, Fan, Fang, and Markussen (2015) provide evidence from Norwegian registry data that mother's employment during children's early years has a differentially adverse effect on the educational attainment of sons relative to daughters. They hypothesize that rising female employment may in part explain the reversal of the male-favorable gap in educational attainment.[7] Aucejo and James (2017) investigate gender gaps in high school graduation and college attendance within race groups, simultaneously examining the role of early childhood behavioral issues, family background, and academic ability. Echoing the results in this paper, they find a substantial contribution of early childhood behavioral outcomes to the gender gap in high school completion. In addition, they find that black boys are more responsive to family background than are black girls.

[6] Lundberg and Pollak (2007), Cherlin (2009), and Cherlin (2014) provide broader discussions of the economics of family structure and its evolution over recent decades. A large sociological literature documents gender differences in outcomes during childhood (cf. Sampson, Morenoff, and Raudenbush 2005; Farrell et al. 2005).

[7] Baker, Gruber, and Milligan (2008) document that the introduction of highly subsidized, universally accessible child care in Quebec during the late 1990s increased maternal labor supply, adversely impacted early childhood developmental outcomes among affected cohorts of children, and generated increased parental time spent with boys relative to girls (Baker and Milligan 2013).

A second strand of literature to which we contribute studies the causal effects of environmental and maternal stresses on children's in utero mortality, birthweight, health, cognitive development, and educational and labor market outcomes. Almond and Currie (2011) and Aizer and Currie (2014) review the large literature on the effect of maternal disadvantage on infant birthweight and infant health, while Almond, Chay, and Lee (2005), Black, Devereux, and Salvanes (2007), and Figlio et al. (2014) document the robust and enduring adverse effect of low birthweight on children's IQ scores, educational attainment, and adult earnings. Several papers confirm the long-standing Trivers-Willard hypothesis (Trivers and Willard 1973) that in utero stress increases the mortality odds of male relative to female fetuses (Norberg 2004, Almond and Edlund 2007). Studies that test for impacts of maternal malnutrition and maternal nutrition supplementation on subsequent academic and labor market outcomes of children who were potentially exposed in utero, find robust, though generally modest, impacts on adult health, skills acquisition, and labor market outcomes (Field, Robles, and Torero 2009; Maccini and Yang 2009; Almond and Mazumder 2011; Almond, Mazumder, and van Ewijk 2015; Greve, Schultz-Nielsen, and Tekin 2015; Chong et al. 2016; Hoynes, Schanzenbach, and Almond 2016). These studies do not, however, find consistent differential impacts by gender on postnatal outcomes. Moreover, where differences are detected, they largely suggest that prenatal disadvantage has a slightly more adverse impact on girls than boys.[8] Assuming plausibly that prenatal and postnatal disadvantage are positively correlated, this (modest) prenatal differential would work *against* our finding that boys are differentially adversely affected by family disadvantage.

A final related strand of work analyzes the causal effects of neighborhoods on developmental, educational, and labor market outcomes, abstracting from the direct effect of family characteristics that are often tightly correlated with neighborhood attributes (e.g., income, race, and educational attainment). Recent work that pairs experimental (Chetty et al. 2016a) or quasi-experimental (Chetty and Hendren 2018a, b) research designs with high resolution tax register data finds that early and prolonged childhood exposure to low or high quality neighborhoods affects labor force participation, earnings, and education in adulthood.

Of particular relevance to our work, two quasi-experimental studies that obtain identification from millions of neighborhood moves, Chetty and Hendren (2018a) and Chetty et al. (2016b), find that boys' outcomes vary more across areas than girls, and that boys' outcomes—particularly employment rates in early adulthood—are differentially sensitive to neighborhood exposure. Our paper confirms the pattern reported by Chetty and Hendren (2018a) and Chetty et al. (2016b) and expands it along multiple dimensions. Most centrally, we demonstrate that although neighborhood environment and family environment are highly correlated, they appear to make substantial independent contributions to the gender gap in behavioral and educational outcomes. Accounting parametrically or non-parametrically for the

[8] See, for example, Greve, Schultz-Nielsen, and Tekin (2015) on the effects of fetal exposure to Ramadan; Hoynes, Schanzenbach, and Almond (2016) on improvements in *in utero* nutrition; Maccini and Yang (2009) on weather shocks around the time of birth in Indonesia; and Field, Robles, and Torero (2009) on prenatal iodine supplementation during the first trimester of pregnancy.

county-level community quality differences documented by Chetty and Hendren (2018b) and Chetty et al. (2016b), or for zip code level measures of neighborhood quality, we find that the bulk of the SES gradient in the gender gap remains—even while schools and neighborhoods also have independent differential effects by gender. This leads to our broader conclusion that diminished child-rearing environments—whether at the household, school, or neighborhood level—appear differentially pernicious for boys.[9]

I. Data and Descriptive Statistics

A. *Principal Data Sources*

Data for our main analysis are drawn from two sources: the universe of birth certificates for the state of Florida for years 1992 through 2002, sourced from the Florida Bureau of Vital Statistics; and linked school records from the 1995–1996 through the 2012–2013 school years from the Florida Department of Education Data Warehouse for children in these birth cohorts who attended the Florida public schools.[10] The Florida Departments of Education and Health matched birth certificate data to schooling records for the purpose of this project using first and last names, exact date of birth and social security number. Of the 2,047,663 births recorded by the Florida Bureau of Vital Statistics during 1992–2002, 1,652,333 were observed in Florida public school data, representing an 80.7 percent match rate. This match rate is almost identical to the percentage of children who are born in Florida, reside there until school-going age, and attend public school, as computed from data from the decennial Census and American Community Survey for years 2000 through 2009.[11]

Florida birth certificates enumerate demographic characteristics of the mother (including age, education, race/ethnicity, marital status, and information whether birth was paid for by Medicaid) as well as health and demographic characteristics of the newborn. We assign children the race/ethnicity of their mother, which is reported on the birth certificate.[12] We use the birth certificate data to create four proxies of family environment: mother's education at birth; mother's age at birth; birth paid

[9] In a paper written subsequent to the current study, Chetty et al. (2018) analyzes the associations among family characteristics, neighborhood, and intergenerational mobility by race and gender. Though intergenerational mobility is distinct from the developmental outcomes analyzed here, Chetty et al. (2018) finds that father presence is a strong predictor of mobility for black boys, both at the family and neighborhood level (though the family-level effect is mediated by income).

[10] The bulk of our analysis is limited to children born between 1994 and 2002, as we only observe neighborhood measures, a proxy for income (Medicaid indicator), and sibship status for children in those birth cohorts. We make use of data from 1992 and 1993 in order to study high school graduation, since our most recent data extract covers outcomes through the high school classes of 2012 and 2013.

[11] Figlio et al. (2013) and Figlio et al. (2014) provide further details on the matching process as well as extensive validity checks on the matching procedure.

[12] This assignment abstracts from mixed race/ethnicity families. If the father does not appear on the birth certificate, we do not observe the race/ethnicity of the father. Evidence from Arcidiacono et al. (2015) indicates that mother race is more predictive of differences in outcomes between black and white male children than is father race. We also analyze the correlation between our assignment of child race/ethnicity and the race/ethnicity of the child later reported in Florida public school records, and find that 90 percent of children born to white mothers report being white in school records; 96 percent of children born to black mothers report being black; and 78 percent of children born to Hispanic mothers report being Hispanic, while 13 percent report being white.

Table 1—Construction of Principal Components SES Index for Family Disadvantage

	First component	Second component
Mother's years of education	0.49	0.71
Married	0.50	−0.67
Non-Medicaid birth	0.52	−0.17
Mother's age at birth	0.49	0.15
Eigenvalue	2.26	0.64
Summary statistics for the first component		
Mean	0.00	
Standard deviation	(1.51)	
Mean white	0.37	
Standard deviation white	(1.41)	
Mean black	−1.06	
Standard deviation black	(1.31)	
Mean Hispanic	−0.27	
Standard deviation Hispanic	(1.51)	

Notes: This table reports the results of a principal components analysis of mother's education (in years), mother's age at birth, non-Medicaid birth indicator, and 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 as well as the white, black, and Hispanic subsamples.

by Medicaid, which we use as an indicator of poverty status at the time of birth; and marital status at the time of birth.[13] We also construct a single composite SES measure based on a principal components analysis of our four proxies of family environment. Details on the construction of this measure are documented in Table 1.

School records include third through eighth grade Florida Comprehensive Assessment Test (FCAT) math and reading scores, as well as records of absence rates and suspensions during each academic year. In addition, for some cohorts of children, the Florida Department of Education recorded the results of a universally administered kindergarten readiness assessment.[14] Each child in the sample can contribute multiple observations, one per each year observed in each grade. Following Autor et al. (2016b), we use school-level gain scores calculated by the Florida Department of Education—schools' average contribution to student outcomes—to measure the quality of elementary and middle schools that children attend. For each school, we compute a simple average of the observed gain scores between 2002 and 2013, which we then convert into a percentile rank in the observed gains distribution across Florida schools.[15] For each student, we construct the cumulative quality of

[13] We include maternal age as a proxy for family environment based on evidence that teen pregnancies are often to girls who are economically disadvantaged (Kearney and Levine 2012; Bertrand and Pan 2013).

[14] The Florida Department of Education recorded kindergarten readiness measures for entering kindergarteners in two sets of kindergarten cohorts: those who entered kindergarten in fall 2001 or before, and those who entered kindergarten in fall 2006 or later. The birth cohorts between 1994 and 2002 who took the kindergarten readiness assessment, therefore, are those born between 1994 and 1996 and those born between 2000 and 2002.

[15] We average the three gain score measures consistently produced over the entire period: percent making gains in reading, percent making gains in math, and percent of bottom quartile students making gains in reading. These scores are available for download at schoolgrades.fldoe.org.

schools attended from grade three through eight, equal to a years-weighted average of the rank quality of all schools attended to that point.

We use two sets of variables to proxy for neighborhood quality: the median income of the zip code of residence observed at birth, computed from the 2000 Census; and a measure of the causal effects of place of residence on economic mobility by Florida birth county from Chetty and Hendren (2018b). The Chetty-Hendren mobility measure corresponds to the estimated percentage gain (or loss) in income at age 26 from spending one more year of childhood in each county in the United States. We extract the causal exposure effects as percentage gains (or losses) in income at age 26 relative to the national mean for parents at the twenty-fifth and seventy-fifth percentile of the national household income distribution. We then average the two values to obtain one treatment effect for each county. Due to linearity in ranks, this measure is the same as the treatment effect on the median person.[16]

B. *Sample Selection*

Table 2 presents summary statistics for the universe of Florida children born to white non-Hispanic, black non-Hispanic, and Hispanic mothers for the years 1994 through 2002 (column 1), which for compactness we refer to as white, black, and Hispanic for the remainder of the paper.[17] The second column drops the 4 percent of records that are missing key variables. The third column includes the 81 percent of column 2 records that were matched to Florida public school records; those not matched to school records have either left the state of Florida or attended private school in the state. Column 4 contains the subset of column 3 records with a valid third-grade test score, while column 5 presents the subset of column 4 records with matched siblings.

Relative to the population of Florida births (column 1), limiting the sample to birth records with complete data (column 2) has almost no effect on birth demographics. Restricting further to births that subsequently appear in Florida public school records (column 3) and eventually obtain a third grade test score (column 4) increases the share of mothers who are black, younger, less educated, and unmarried at the time of childbirth. These compositional changes are consistent with the greater cross-state mobility of high SES adults and higher private school attendance rates of their children (Molloy, Smith, and Wozniak 2011). While the matched sample represents a more disadvantaged population than the full sample, the gender composition of those matched to school records does not differ appreciably from the full population of births, and there are virtually no differences in birthweight

[16] We thank Jeremy Majerovitz for expert assistance with the Chetty-Hendren data. We use data from their Online Data Table 2 (available at http://equality-of-opportunity.org/index.php/data): Preferred Estimates of Causal Place Effects by County.

[17] We exclude the 2.7 percent of births where mothers do not identify as white, black, or Hispanic. We also exclude the 24 percent of births that are to immigrant mothers since these families are too heterogeneous to defy any simple characterization. As reported in Autor et al. (2016a), gender gaps among children born to immigrant mothers fall in between the black and Hispanic gaps, and the proportion of the gap that is explained by the differential effect of family SES on boys is similar to that of the Hispanic population.

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

	White, black, and Hispanic births (1)	With complete data (2)	Matched to Florida school records (3)	Matched to third grade test score (4)	Sibling sample with third grade test score (5)
White non-Hispanic	67.7	67.6	65.1	64.0	59.2
Black non-Hispanic	21.7	21.7	24.1	25.1	29.5
Hispanic	10.6	10.7	10.8	10.9	11.3
High school dropout	18.7	18.6	20.4	21.0	22.7
High school graduate	60.1	60.3	61.8	62.3	59.4
College graduate	20.9	21.1	17.8	16.7	17.9
Age 21 or below	24.1	24.0	26.1	26.8	27.9
Age between 22 and 29	41.9	41.9	42.0	42.0	44.8
Age between 30 and 35	24.9	25.0	23.4	23.0	21.5
Age 36 or above	9.1	9.1	8.5	8.2	5.8
Married at time of birth	63.5	63.7	60.5	59.2	59.4
Medicaid-funded birth	43.9	43.6	47.2	48.2	49.9
Boy	51.3	51.3	51.0	50.5	50.8
Birth weight (grams)	3,337	3,337	3,324	3,322	3,325
Observations	1,209,870	1,157,930	935,993	770,384	301,298

Notes: This table reports summary statistics for the Florida statewide data. The first column is the full sample of singleton Florida births 1994–2002, excluding immigrants. The second column drops the 4 percent of records that are missing key variables. The third column contains the 81 percent of column 2 records that were matched to Florida school records. 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 test score; and the fifth column is the subset of column 4 in the matched sibling sample—children in families with two or more children born 1994–2002. All demographic characteristics are derived from the birth certificate. White, black, and Hispanic are mutually exclusive categories. The first 12 rows present maternal characteristics, while the last 2 rows present child characteristics.

between the full population and the matched sample.[18] When we further limit the sample to matched sibling births (column 5), we observe relatively fewer white mothers, slightly more college educated mothers, and slightly more births funded by Medicaid.

Table 3 summarizes key demographic characteristics for our main sample observed either at birth or during K–12 schooling according to the race and ethnicity of the mother. These include mother's age and education, whether the birth was paid for by Medicaid, marital status at birth, the median income in the zip code of the mother's residence at the time of birth, the Chetty-Hendren (2018b) measure of county economic mobility associated with that zip code, and the rated quality of Florida public schools subsequently attended. The pronounced contrasts among race and ethnic groups in each of these measures highlights the degree to which family disadvantage differs systematically across these broad demographic groups. For example, the percentage of children born to unmarried mothers ranges from 27.2 among whites to 75.5 among blacks. Comparisons of maternal education, Medicaid-funded births, and zip code income reveal similarly stark contrasts, underscoring that comparisons across race and ethnic groups are also implicitly comparisons across education groups, income levels, and family types.

[18] The slight reduction in the fraction of male births (from 51.3 percent to 50.5 percent) as the sample is restricted to more disadvantaged mothers is consistent with the Trivers-Willard (1973) hypothesis.

Table 3—Descriptive Statistics: Selected Characteristics, by Race and Ethnicity

	All (1)	White non-Hispanic (2)	Black non-Hispanic (3)	Hispanic (4)
Married	59.1	72.8	24.5	58.5
Mother high school dropout	21.1	16.3	30.9	26.8
Mother high school grad	62.2	62.6	62.8	58.9
Mother college grad	16.6	21.1	6.3	14.3
Non-Medicaid birth	51.6	63.2	24.5	46.1
Mother age at birth	26.24	27.39	23.84	25.09
Median zip code income	44,823	47,874	37,251	44,439
	(13,289)	(13,123)	(10,458)	(13,057)
Observations	796,701	509,372	200,812	86,517
Chetty/Hendren CZ mobility measure	−0.24	−0.24	−0.25	−0.27
	(0.20)	(0.20)	(0.19)	(0.19)
Observations	795,980	508,838	200,672	86,470
School quality	51.0	54.1	42.4	52.8
	(14.8)	(14.5)	(12.4)	(13.7)
Observations	764,684	487,358	195,110	82,216

Notes: This table reports summary statistics for the sample of non-twin singletons born 1994–2002, by race and ethnicity of the mother. Standard deviations for continuous measures are given in parentheses. The sample is reduced for the Chetty/Hendren CZ mobility and the school quality variables due to the lack of availability of these measures for a subset of Florida counties and schools.

C. *Gender Gaps by Family and Maternal Characteristics*

Figure 1 sets the stage for our empirical inquiry by plotting gender contrasts in educational and behavioral outcomes by family demographics along four dimensions: kindergarten readiness (a behavioral and academic measure); school attendance rates (a behavioral measure); standardized reading scores (an academic measure); and on-time high school graduation rates (a labor market measure). Kindergarten readiness and high school graduation are observed for a subset of cohorts. Appendix Tables A1 and A2 provide the corresponding summary statistics. Attendance rates (the complement of absence rates) and reading scores are observed during grades three through eight for 1994 to 2002 cohorts; kindergarten readiness is observed for cohorts born between 1994 and 1996 and those born between 2000 and 2002; and high school completion is observed for the 1992 and 1993 cohorts, which had reached the age of on-time high school completion by the end of our sample.

The first panel of the figure highlights the cross-race and cross-ethnicity differences in the gender gap in academic and behavioral outcomes. Among children born to white mothers, the boy-girl gap in the attendance rate is −0.17 percentage points, the boy-girl gap in elementary and middle school reading scores is −0.14 standard deviations (SDs), and the boy-girl gap in both kindergarten readiness as well as on-time high school graduation is about −6 percentage points (i.e., girls start kindergarten more ready and also have higher graduation rates). The gap in each of these outcomes is monotonically widening (becoming less favorable to boys) as we move the focus of

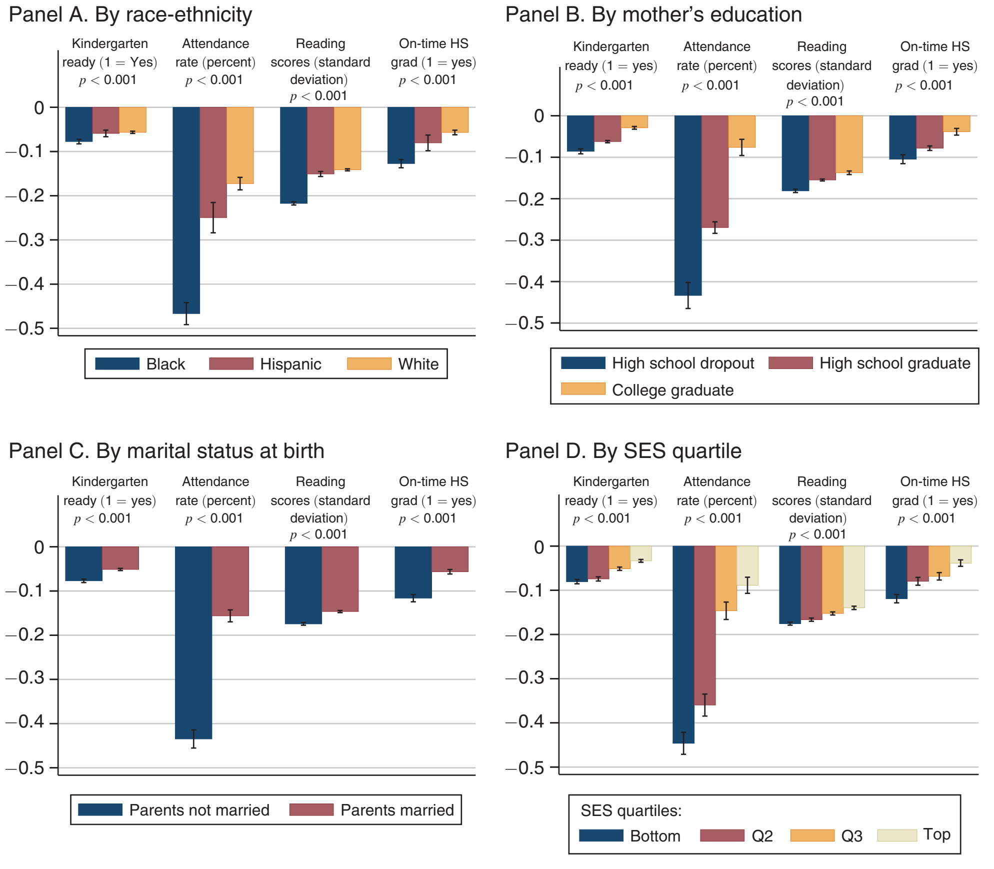

Figure 1. Boy-Girl Gaps in Kindergarten Readiness, Attendance, Reading Scores, and On-Time High School Completion by Family Characteristics

Notes: This figure plots gender gaps in kindergarten readiness rates, attendance rates, FCAT reading scores, and on-time high school completion rates, by gender and family demographics. Kindergarten readiness rates are computed from the pooled sample of non-twin singletons born in cohorts 1994 to 1996 and 2000 to 2002. Attendance rates and FCAT reading scores are computed from the pooled sample of non-twin singletons spanning grades three through eight born between 1994 and 2002. On-time high school completion indicates a high school diploma within four years of entering and is computed from the pooled sample of non-twin singletons who were observed in Florida public schools until at least ninth grade. High school completion rates are computed from the 1992 and 1993 birth cohorts. The SES measure for kindergarten readiness, attendance, and reading scores is constructed as the first component of a principal components analysis of years of maternal education, maternal age, an indicator for non-Medicaid funded birth, and an indicator for parents married at the time of birth. The SES measure for high school is constructed as the first component of a principal components analysis of years of maternal education, maternal age, and an indicator for parents married at the time of birth. Ninety-five percent confidence intervals are represented by whiskers on each bar.

comparison from whites to Hispanics to blacks. Among children of black mothers, the boy-girl gaps in attendance, reading scores, and on-time graduations are, respectively, -0.47 percentage points, -0.22 SDs, and -12.7 percentage points—in each case, two to three times as large as among whites.[19] The boy-girl gap in kindergarten readiness

[19] The gender gap in standardized math scores has the same monotonic pattern, but there is a male-favorable math score gap in advantaged families and a female-favorable math score gap in disadvantaged families.

among blacks is −7.8 percentage points, nearly one and a third the size of the boy-girl gap among white children. The gender gaps in outcomes among children of Hispanics mothers fall roughly in between those of whites and blacks on these measures.

Panels B through D of Figure 1 plot analogous contrasts by maternal education (high school dropout, high school graduate or some college, four year college degree or higher), family structure at birth (married, not married), and SES quartile.[20] The boy-girl deficits in behavioral and educational outcomes are robustly larger in families with greater disadvantage, whether measured by maternal education, family structure, or the SES composite.

II. Empirical Framework

Our empirical objective is to isolate the effect of family disadvantage on the gender gap in childhood behavioral and educational outcomes. Since family types are not randomly assigned to children, the below framework formalizes how the intrinsic correlation between parents and children, operating through both environmental and hereditary channels, potentially confounds the effect of family environment. We then discuss how our estimation strategy seeks to address these confounds.

Let $Y_b \in \{Y_b^0, Y_b^1\}$ and $Y_g \in \{Y_g^0, Y_g^1\}$ equal the potential outcomes of boys and girls, b and g, born into a family whose socioeconomic status at the time of the children's birth is $D \in \{0,1\}$. For expositional simplicity, we treat D as discrete (i.e., a family is either disadvantaged or advantaged), but we relax this assumption in the estimation. We observe only one set of potential outcomes as a function of family disadvantage, D:

$$(1) \quad Y_b = Y_b^1 \times D + Y_b^0 \times (1 - D) \quad \text{and} \quad Y_g = Y_g^1 \times D + Y_g^0 \times (1 - D).$$

We expect that potential outcomes will differ between boys and girls,

$$(2) \quad E\left[Y_b^1\right] \neq E\left[Y_g^1\right], \quad E\left[Y_b^0\right] \neq E\left[Y_g^0\right],$$

and further, that potential outcomes will not be independent of family disadvantage for either boys or girls,

$$(3) \quad E\left[Y_s^1 | D = 1\right] \neq E\left[Y_s^1 | D = 0\right],$$

$$E\left[Y_s^0 | D = 1\right] \neq E\left[Y_s^0 | D = 0\right] \quad \text{for } s \in \{b, g\}.$$

[20] Due to data limitations, we construct a separate SES index for the cohorts utilized for high school completion outcomes. This SES index is based on maternal age, education, and marital status at the time of birth.

These confounds mean that absent random assignment of disadvantage to families, simple between-family contrasts will not provide a valid estimate of the effect of family disadvantage on outcomes Y for either boys or girls. To see why, consider the non-experimental contrast between the outcomes of boys born to disadvantaged versus advantaged families. This contrast is

$$(4)\quad E\left[Y_b|D = 1\right] - E\left[Y_b|D = 0\right] = \left\{E\left[Y_b^1|D = 1\right] - E\left[Y_b^0|D = 1\right]\right\} + \left\{E\left[Y_b^0|D = 1\right] - E\left[Y_b^0|D = 0\right]\right\},$$

where the first bracketed term on the right-hand side is the average causal effect of family disadvantage on outcome Y for disadvantaged boys, and the second bracketed expression is a bias term, stemming from differences in potential outcomes between boys born to disadvantaged versus advantaged families. Both intuition and data suggest that this bias term will be nonzero (per equation (3)): children born to advantaged families are likely to have genetic and health advantages at birth that may yield more favorable outcomes, holding child-rearing circumstances constant.

In this paper, we select as our outcome of interest the difference in the gender *gap* in outcomes, $\hat{Y}^1 \equiv Y_b^1 - Y_g^1$ and $\hat{Y}^0 \equiv Y_b^0 - Y_g^0$, and contrast the gender gap between advantaged and disadvantaged families. Our hypothesis is that developmental outcomes of boys are more elastic than those of girls to the quality and quantity of family inputs—hence, the same exposure to family advantage or disadvantage affects boys and girls differently.

Under what conditions can we identify the causal effect of family disadvantage on the gender gap? Our key identifying assumption is that the *latent* gap in childhood outcomes between boys and girls is as good as randomly assigned to family types, that is, $\hat{Y}^1, \hat{Y}^0 \perp D$. Under this assumption, *observed* variation in the gender gap that differs systematically across family types will reflect the causal effect of family advantage on the gender gap (rather than a reification of the latent gap).

This assumption permits identification of the causal effect of family disadvantage on the gender gap accordingly by contrasting gender differences in outcomes across disadvantaged and advantaged families:

$$(5)\quad \begin{aligned} & E\left[\hat{Y}^1|D = 1\right] - E\left[\hat{Y}^0|D = 0\right] \\ &= E\left[Y_b^1 - Y_g^1|D = 1\right] - E\left[Y_b^0 - Y_g^0|D = 0\right] \\ &= E\left[Y_b^1 - Y_g^1|D = 1\right] - E\left[Y_b^0 - Y_g^0|D = 1\right], \end{aligned}$$

where the second equality follows from our assumption above. This double-difference eliminates the dependency between family disadvantage and potential outcomes by contrasting boys and girls of the same family background; it further eliminates the dependency between gender and potential outcomes (arising from intrinsic gender differences) by contrasting the gender gap across advantaged and disadvantaged families.

Is this assumption plausible? Suppose, for example, in violation of our identifying assumption, that family SES differentially affected the fetal development of sons relative to daughters—which could occur if the male fetus is more sensitive than the female fetus to maternal stress levels during pregnancy. In this case, the contrast formed in equation (5) would confound any causal effect of family disadvantage on the gap in outcomes between boys and girls with the differential in utero impact of maternal stress on the subsequent development of children of each sex. We explore this possibility by assessing whether gender gaps in neonatal health, measured by birthweight, maternal health, and a composite health index, differ across families of varying socioeconomic levels. As documented in Section IV, we find little evidence that these gaps vary substantially or systematically with family disadvantage.[21]

There are two additional sources of potential bias in our estimation of the effect of family disadvantage on the gender gap. The first occurs if family structure is endogenous to the sex of children, as suggested by Dahl and Moretti (2008), leading to the possibility that the potential outcomes of girls and boys may be differentially correlated with family disadvantage. In our estimation, we attempt to address this confound by providing a subset of our analyses that limit our sample to families with two or more children, thus sweeping out unobserved family attributes by forming within-family, cross-gender comparisons. The second source of bias stems from the relationship between family SES and other environmental factors, such as attributes of schools and neighborhoods. Since the effect of SES on children's outcomes will incorporate these multiple channels of influence, our empirical analysis will further explore the robustness of our family SES estimates to the inclusion of detailed school quality and neighborhood controls.

A. *Implementation*

We use a simple regression framework to quantify our motivating observation that there is a substantially greater female advantage in behavioral and early academic preparedness among children of black and Hispanic households relative to children of white households. This observed gap serves as a benchmark for assessing the explanatory power of our subsequent models. To obtain this baseline, we estimate the following regression:

$$(6) \qquad Y_i = \alpha + \beta_1 Boy_i + \beta_2(Boy_i \times Black_i) + \beta_3(Boy_i \times Hispanic_i) + \mathbf{X}'_{1,i}\lambda_1 + \mathbf{X}'_{2,j(i)}\lambda_2 + e_i,$$

where Y_i represents an outcome for child i, Boy_i is an indicator variable for whether the child is male, and $Black_i$ and $Hispanic_i$ are indicators for whether the mother of child i belongs to one of those mutually exclusive race or ethnic categories, with white, non-Hispanic mothers serving as the reference category. The vector $\mathbf{X}'_{1,i}$ controls for time-invariant child attributes, including birth order and month and year of birth. The vector $\mathbf{X}'_{2,j(i)}$ controls for the main effects of maternal and family

[21] Furthermore, the studies cited in the introduction assessing the impact of *in utero* shocks on postnatal outcomes generally find that adverse impacts are slightly more pronounced for girls than boys, which would work against a finding that postnatal family disadvantage is differentially detrimental to boys.

environment characteristics, including mother j's education, Medicaid receipt, maternal age, and marital status. We denote the dependence between mother and child via the subscript $j(i)$ because a subset of models will include mother fixed effects to form within-family, cross-sibling contrasts.

In this initial descriptive regression model, the coefficient β_1 measures the gender difference in outcome Y_i for the reference category of children of white, non-Hispanic mothers. The coefficients β_2 and β_3 correspond to the demographic differentials of interest, indicating how the gender gap varies with the race-ethnicity status of the mother.

In a second step, we apply our proposed identification strategy to estimate the effect of family disadvantage on the gender gap in academic and behavioral outcomes by augmenting the baseline model to permit the effect of family disadvantage D_i to differ by child gender:

$$(7) \qquad Y_i = \alpha + \beta_1' Boy_i + \beta_2'(Boy_i \times Black_i) + \beta_3'(Boy_i \times Hispanic_i) + \beta_4'(Boy_i \times D_i) + \mathbf{X}_{1,i}'\lambda_1' + \mathbf{X}_{2,j(i)}'\lambda_2' + e_i.$$

Here, D_i is either a vector of proxies for family environment or a composite measure of family disadvantage. The coefficient β_4' on the interaction term $(Boy_i \times D_i)$ captures the differential responsiveness of boys to family disadvantage. The interpretation of β_4' as the *causal* effect of family disadvantage on the gender gap in cognitive and behavioral outcomes hinges critically on the unconfoundedness assumption. We assess the plausibility of this assumption in Section IV by estimating equation (7) for pregnancy health and at-birth child health outcomes. Because our research design is non-experimental, the reader may prefer to view these relationships as associations rather than causal effects, though we believe the evidence below on unconfoundedness supports the latter (i.e., stronger) interpretation.

Through a comparison of the coefficients from the model from equation (6) and the augmented model from equation (7), we infer what fraction of the race-ethnicity gradient in the gender gap is proximately explained by the differential effect of family disadvantage on boys. Specifically, we compare β_2 to β_2' and β_3 to β_3'. Take, for example, the comparison of the black-white gender gap in the baseline model relative to the augmented model, as estimated by β_2 and β_2', respectively. If we find that the black-white gender gap declines when we permit family disadvantage to differentially affect boys, i.e., β_2' is smaller in (absolute) magnitude than β_2, we would attribute this to the fact that disadvantage exerts a disproportionate negative effect on boys relative to girls and, critically, is more prevalent among black than white families.

As described in the Empirical Framework, if family disadvantage is endogenous to child gender, the condition that boys and girls are, on average, exposed to the same family environment would be violated. In order to test for this confound, we additionally implement a within-family, cross-sibling comparison, where we limit our sample to families with two or more siblings and include mother fixed effects $\delta_{j(i)}$ to control for unobserved family-specific factors. In practice, these two

Table 4A—Family Disadvantage and the Gender Gap in Behavioral Outcomes

	Panel A. Absence rate (percent)				*Panel B. Suspension rate (percent)*			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Boy × Mother years of education		−0.02				−0.75		
		(0.01)				(0.03)		
Boy × Married		−0.12				−1.28		
		(0.03)				(0.13)		
Boy × Non-Medicaid birth		−0.19				−3.41		
		(0.03)				(0.13)		
Boy × Mother age at birth		0.01				−0.03		
		(0.00)				(0.01)		
Boy × SES index			−0.07	−0.08			−2.17	−2.40
			(0.01)	(0.01)			(0.03)	(0.07)
Boy	0.19	0.43	0.21	0.24	8.07	21.75	8.58	7.81
	(0.01)	(0.07)	(0.01)	(0.02)	(0.05)	(0.33)	(0.05)	(0.12)
Boy × Black	0.27	0.15	0.17	0.20	4.50	1.67	1.48	2.31
	(0.02)	(0.03)	(0.03)	(0.04)	(0.13)	(0.15)	(0.14)	(0.30)
Boy × Hispanic	0.07	0.03	0.03	0.08	−0.46	−1.73	−1.80	−1.61
	(0.03)	(0.03)	(0.03)	(0.04)	(0.14)	(0.14)	(0.14)	(0.30)
Sibling fixed effects	No	No	No	Yes	No	No	No	Yes
Mean of Y		5.11		4.89		12.82		12.37
Number of children		792,729		301,128		792,729		301,128

Notes: This table reports the results of regression models where the dependent variables are the absence rate (panel A) and the suspension rate (panel B) from grades three through eight. Columns 1–3 use the sample of non-twin singletons born 1994–2002. Column 4 restricts the sample to matched siblings—children in families with two or more births between 1994 and 2002. All columns include controls for child year and month of birth, maternal race-ethnicity, and birth order within family. Columns 1 and 2 further control for maternal education, maternal age, Medicaid funded birth, and marital status at birth. Columns 3 and 4 control for the main effect of the SES index. The SES index is constructed as the first component of a principal components analysis of years of maternal education, maternal age, non-Medicaid funded birth indicator, and marital status at birth. Column 4 includes mother fixed effects. Standard errors are clustered at the child level in columns 1–3 and mother level in column 4. See online Appendix Table O1 for an expanded set of results.

specifications (all children and siblings with mother fixed effects) yield comparable results.

Equation (7) imposes the restriction that the impact of disadvantage on the gender gap in outcomes is constant across race and ethnicity groups—that is, the coefficient on $Boy_i \times D_i$ does not require a race-ethnicity subscript. If this restriction does not hold, the precise contribution of disadvantage to cross-group variation in the gender gap may be ambiguous for the simple reason that the share of the cross-group gap explained by disadvantage will depend upon which group-specific $Boy_i \times D_i$ slope is used for the calculation. We test this restriction of constant slopes in Section IIIC and we elaborate on the results below.

III. Main Results: Gender Gaps in Behavioral and Academic Outcomes

Our main analysis estimates the effect of family disadvantage on the gender gap in two behavioral outcomes (absence and suspension rates) and two academic outcomes (standardized reading and math test scores) that are observed in elementary through middle school, as well as for kindergarten readiness. We subsequently assess whether these effects on medium-term outcomes carry forward to a directly market-relevant outcome: high school completion.

A. *Elementary and Middle School Behavioral and Academic Outcomes*

Behavioral Outcomes: Absences and Suspensions.—The first column in panel A of Table 4A reports estimates of equation (6) for the gender gap in absence rates, by mother's race-ethnicity. The conditioning variables for this model include mother's age, race/ethnicity, and education, child's birth order, birth month, and birth year, and indicators for Medicaid-paid birth and mother's marital status at the time of delivery. The coefficient on the boy main effect, corresponding to the conditional mean difference between white, non-Hispanic boys and girls, indicates that boys from this ethnic group have grade three through eight absence rates that average 0.19 percentage points higher than those of girls. Confirming the qualitative patterns in attendance in Figure 1, the coefficients on the interactions between the boy main effect and each of the race-ethnicity main effects reveal that the boy-girl gap in absence rates is higher by 0.27 and 0.07 percentage points for black and Hispanic children, respectively. We refer to these interactions as *excess* gender gaps, denoting the unexplained additional gender gaps detected among black and Hispanic children relative to white children from families with the same maternal education, maternal age, Medicaid status, and marital status.

Column 2 augments this model with interactions between the boy indicator and variables that proxy for family advantage at birth—marital status, poverty, and maternal education and age, as per equation (7). Consistent with our primary hypothesis, the interaction terms between the boy dummy and various measures of advantage are mostly negative, indicating that the gender gap in absence rates is smaller among children born to more advantaged families. Relative to children born to non-married parents, the boy-girl gap in absences is 0.12 percentage points lower where the parents were married at the child's birth. Higher levels of family income predict an additional reduction in this gap of 0.19 percentage points, while each additional year of maternal education reduces the gap by 0.02 percentage points. The coefficient on the interaction of maternal age and boy is positive, which is inconsistent with our predictions for family advantage reducing the gender gap in behavioral outcomes, though this effect is quantitatively quite small. Noting that these effects are additive, the model implies that the boy-girl gap in school absences is 0.39 percentage points lower among children born to married parents with college-educated mothers (i.e., 16 years of education) whose births were not funded by Medicaid relative to those born to non-married parents with mothers with a high school degree (i.e., 12 years of education) who utilized Medicaid payments for birth.

Are effects of this magnitude economically consequential? A useful way to answer this question is to benchmark these effects relative to the excess gender gaps among black and Hispanic children. If the effect of family disadvantage on the gender gap is economically consequential *and* if disadvantage is systematically greater among minorities than whites, inclusion of the interactions between gender and disadvantage will reduce these excess gender gaps. Column 2 confirms this conjecture. Simply including interactions between gender and family structure, Medicaid

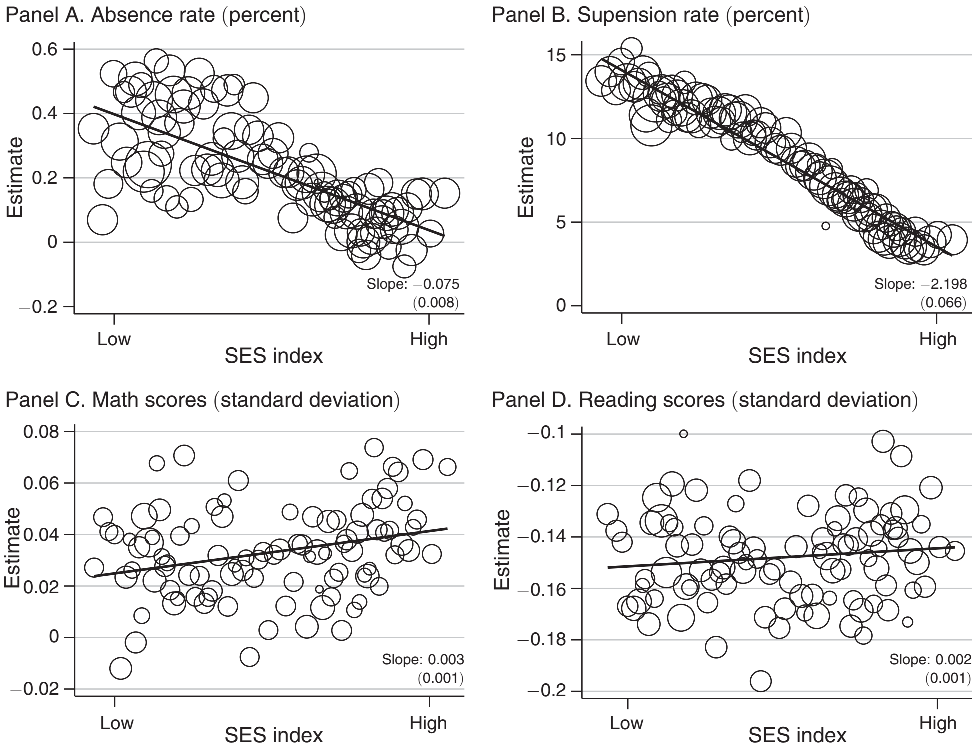

Figure 2. Boy-Girl Gap in Behavioral and Academic Outcomes by Family SES at Birth

Notes: This figure plots regression-adjusted mean absence rates, suspension rates, and FCAT math and reading scores against values of the SES index. Graphs suppress bins with very small sample sizes for expositional clarity. Lines come from OLS regressions fitted through all bins and weighted by bin size. The sample is non-twin singletons born 1994–2002. The regression controls include child gender, year and month of birth, maternal age at birth, birth order within family, the main effect of the SES index, and the race-ethnicity of the mother and their interactions with child gender. The SES measure is constructed as the first component of a principal components analysis of years of maternal education, maternal age, non-Medicaid birth indicator, and indicator for parents married at the time of birth.

indicator, and maternal age and education reduces the excess black gender gap by almost half, from 0.27 to 0.15 percentage points, with larger proportional declines (from a smaller base) among Hispanic children. Thus, the bulk of the black-white excess gender gap is proximately explained (in the sense of our decomposition above) by the disparate effect of family disadvantage on boys relative to girls. If our causal framework above is valid—a subject we return to below—we can infer that nearly half of the excess gender gap in absenteeism among black children is due to the differential adverse effect of disadvantage on boys relative to girls rather than factors specific to black children per se. The role that disadvantage plays in the excess gender gaps for Hispanic families is equally large in proportional terms.

To streamline exposition, column 3 of the table subsumes the interaction terms in the first four rows of column 2 into a single composite SES measure based on a principal components analysis of marital status, the proxy for poverty, and maternal age and education (in years). Similar to the higher dimensional specification in column 2, the interaction between SES and gender reduces the black excess gender gap in absences from 0.27 to 0.17. This reduction follows directly from two facts: the large SES gap between white and black families $(+1.44\sigma)$ and the estimated differential effect of family disadvantage on boys relative to girls. Figure 2, panel A plots the SES gradient in the gender gap using a bin-scatter of the boy-girl gap in absences against the composite SES index while conditioning on all of the covariates used in the corresponding estimate in Table 4A (column 3 of panel A). This figure underscores the robust SES gradient in the gender gap in absences.[22]

A potential threat to validity is that the family types to which boys and girls are exposed could differ systematically along unobservable dimensions, in violation of our main identification assumption. We offer a robustness check on this assumption in column 4 of Table 4A by testing whether the estimated relationships between family disadvantage and the gender gap continue to hold when we move to a sample of siblings and include mother fixed effects to form within-family, cross-gender contrasts. These estimates reinforce the prior conclusions: the boy-girl gap in absence rates is larger among black and Hispanic families than among white families; conditional on race and ethnicity, the boy-girl gap is larger in lower SES families; and inclusion of the composite SES index substantially reduces these excess gender gaps.[23] These results suggest that the ex ante concerns regarding the endogeneity of family types to child gender do not appear to be consequential in practice.[24]

We perform the analogous estimation exercise for the gender gap in school suspensions in panel B of Table 4A. Approximately 13 percent of Florida public school children are suspended for at least 1 day per school year during grades three through eight (Appendix Table A1). But suspension rates are more than twice as high among boys as girls (17.3 versus 8.3 percent), and the gender gap is over 50 percent larger among blacks than whites (12.6 versus 8.0 percent). The bin-scatter in Figure 2 panel B documents that conditional on race and ethnicity, the gender gap in suspensions is far smaller in families where children are born to married parents, where mothers are older and better educated, and where family income is higher. The second column of panel B of Table 4A quantifies these patterns. Collapsing our multiple indicators of family disadvantage into the composite SES index in column 3 yields strong evidence of an SES gradient in the gender gap: each standard deviation increase in disadvantage (a reduction in the SES index of 1.51) increases the boy-girl gap in suspensions by 2.2 percentage points. Accounting for the differential effect of disadvantage of boys relative to girls reduces the black excess suspension gap by about two-thirds. We draw the same inference when contrasting the gender gap in suspensions among siblings born to the same mother in column 4.

[22] The plot includes point estimates for the larger SES bins only, and also reports the slope of an OLS regression fit to all points. This slope is comparable to the corresponding estimate in Table 4A but does not match exactly given the differences in the specifications.

[23] Note that in the sibling sample, the baseline excess gender gaps differ from those in the singletons sample. See online Appendix Table O3 for an expanded set of estimates using the sample of siblings, which should be used as baseline estimates for comparisons with column 4.

[24] Not all selection concerns are resolved by inclusion of family fixed effects: households with mixed-gender siblings may differ from other households, for example. The fixed effects estimates demonstrate that the SES gradient in the gender gap holds *within* households.

Table 4B—Family Disadvantage and the Gender Gap in Academic Outcomes

	Panel A. Mathematics scores (standard deviation)				*Panel B. Reading scores (standard deviation)*			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Boy × Mother years of education		0.005				0.004		
		(0.001)				(0.001)		
Boy × Married		−0.009				−0.001		
		(0.005)				(0.005)		
Boy × Non-Medicaid birth		0.013				0.012		
		(0.005)				(0.005)		
Boy × Mother age at birth		−0.001				−0.002		
		(0.000)				(0.000)		
Boy × SES index			0.005	0.007			0.002	0.005
			(0.001)	(0.003)			(0.001)	(0.002)
Boy	0.034	−0.013	0.033	0.054	−0.147	−0.163	−0.148	−0.121
	(0.002)	(0.013)	(0.002)	(0.004)	(0.002)	(0.012)	(0.002)	(0.004)
Boy × Black	−0.119	−0.116	−0.112	−0.120	−0.070	−0.068	−0.067	−0.078
	(0.004)	(0.005)	(0.005)	(0.009)	(0.004)	(0.005)	(0.004)	(0.009)
Boy × Hispanic	−0.025	−0.023	−0.023	−0.028	−0.009	−0.008	−0.007	−0.002
	(0.006)	(0.006)	(0.006)	(0.011)	(0.006)	(0.006)	(0.006)	(0.011)
Sibling fixed effects	No	No	No	Yes	No	No	No	Yes
Mean of Y		0.060		0.054		0.080		0.042
Number of children		785,664		297,907		785,673		297,938

Notes: This table reports the results of regression models where the dependent variables are standardized FCAT math scores (panel A) and standardized FCAT reading test scores (panel B) from grades three through eight. Columns 1–3 use the sample of non-twin singletons born 1994–2002. Column 4 restricts the sample to matched siblings—children in families with two or more births between 1994 and 2002. All columns include controls for child year and month of birth, maternal race-ethnicity, and birth order within family. Columns 1 and 2 further control for maternal education, maternal age, non-Medicaid funded birth, and marital status at birth. Columns 3 and 4 control for the main effect of the SES index. The SES index is constructed as the first component of a principal components analysis of years of maternal education, maternal age, non-Medicaid funded birth indicator, and marital status at birth. Column 4 includes mother fixed effects. Standard errors are clustered at the child level in columns 1–3 and mother level in column 4. See online Appendix Table O2 for an expanded set of results.

Academic Outcomes: Test Scores.—We measure academic performance using standardized mathematics and reading tests administered annually during elementary and middle school. Relative to children in white families, black and Hispanic boys perform worse than girls in mathematics and reading. Distinct from the two outcomes analyzed above (absences and suspensions), we find in Table 4B that family disadvantage contributes only modestly to the cross-race and cross-ethnic group variation in the gender gap in these educational measures. The coefficients on some measures of family advantage (maternal education and income) have the expected sign (reducing relative boy-girl disadvantage in math and reading), but the coefficient on maternal age and marital status have the opposite sign. When combined into a composite measure, family SES is a statistically significant predictor, but of modest economic magnitude (also see in the bin-scatter in Figure 2, panels C and D). Accounting for SES differences explains at most 6 and 4 percent of the excess black-white gender gap in math and reading test scores, respectively. The substantially stronger effects for behavioral outcomes are consistent with findings from the Perry Preschool Program, which provided intensive early enrichment to low-income minority children, leading to improved labor market and demographic outcomes in adulthood. The Perry Program did not have lasting impacts on IQ scores, but did augment personality traits that are predictive of labor market success (Heckman and Kautz 2012).

Early Academic and Behavioral Outcomes.—We use the Florida Department of Education's assessments of kindergarten readiness to construct a dichotomous measure of readiness based on the approach described in Figlio et al. (2013). Appendix Table A3 reports the estimates of the effect of family disadvantage on the gender gap in kindergarten readiness. Across all specifications, a one standard deviation increase in the SES index raises the readiness of boys relative to girls—that is, reduces their disadvantage—by more than a full percentage point. Thus, the apparent effect of family disadvantage on the gender gap in schooling outcomes emerges as early as age five. The inclusion of the interaction of boy and SES reduces the excess black-white disparity in kindergarten readiness among boys by 60 to 77 percent depending on specification, and the adjusted differentials are no longer statistically different from zero.

We have also explored how the SES gradient in the gender gap in early outcomes evolves between grades three and eight using successive observations across grades. Both the SES gap and the gender gap in absences, suspensions, and test scores increases year-over-year across grades: low SES children fare increasingly poorly relative to high SES children, and boys fare increasingly poorly relative to girls. We do not, however, find a consistent triple-interaction between these two forms of divergence.

B. *High School Completion*

The cumulative adverse effect of family disadvantage on the boy-girl gap in behavioral and academic outcomes in kindergarten through middle school may contribute to gender gaps in downstream market outcomes, including educational attainment and earnings. We test this hypothesis by analyzing the key market outcome that is available in our data: high school graduation.[25] We ask whether family disadvantage affects the gender gap in high school graduation, and whether the impact of family disadvantage on third through eighth grade behavioral outcomes and test scores can account for this relationship. This analysis is limited to the oldest cohorts in our sample, those born in 1992 and 1993, for whom on-time graduations are presently available.

To assess the relationship between family advantage and the gender gap in high school graduations (summarized in Appendix Table A2), Table 5 presents two OLS models fit to our Florida sample: one controlling for an SES index constructed with the demographic characteristics available for the 1992 and 1993 cohorts (maternal education, maternal age, marital status), the boy $\times$ race-ethnicity dummies, and all child-level controls; and the second augmented with the interaction between

[25] While high school diplomas are not a market outcome per se, high school degrees are priced in the labor market (see Autor (2014) for discussion).

Table 5—Family Disadvantage and the Gender Gap in High School Graduation

	Panel A. On-time high school graduation		*Panel B. 5+ years of high school*		*Panel C. High school dropout*	
	(1)	(2)	(1)	(2)	(1)	(2)
Boy × SES index		1.46		−0.82		−0.64
		(0.18)		(0.14)		(0.16)
Boy	−6.04	−6.45	3.92	4.15	2.12	2.30
	(0.26)	(0.27)	(0.18)	(0.19)	(0.22)	(0.23)
Boy × Black	−6.72	−4.98	2.98	1.99	3.75	2.99
	(0.53)	(0.58)	(0.43)	(0.47)	(0.43)	(0.47)
Boy × Hispanic	−2.10	−1.29	0.34	−0.11	1.75	1.40
	(0.91)	(0.91)	(0.69)	(0.69)	(0.76)	(0.77)
Mean of Y	70.42		12.75		16.83	
Number of children	161,537		161,537		161,537	

Notes: This table reports the results of regression models where the dependent variable is, in turn, on-time graduation, continuation in high school, and dropout, from the sample of non-twin singletons who were born in 1992 or 1993 and observed in Florida public schools until at least ninth grade. On-time high school completion takes on a value of 100 if a student obtains a high school diploma within 4 years of entering, and is 0 otherwise. 5+ years of high school takes on a value of 100 if the student is enrolled in high school more than 4 years after entry but has not yet dropped out, and is 0 otherwise. High school dropout takes on a value of 100 if a student does not earn a high school diploma and is no longer enrolled in high school 5+ years after entry, and is 0 otherwise. The SES index is constructed as the first component of a principal components analysis of years of maternal education, maternal age, and marital status at birth. Robust standard errors are in parentheses.

child gender and the SES index.[26] The high school graduation outcomes shown in Table 5—on-time graduation, delayed graduation, and dropout—are exhaustive and mutually exclusive. The first column of panel A shows that the conditional mean on-time high school graduation rate of white boys is 6.0 percentage points below that of white girls, and that there is an additional (excess) gap of 6.7 and 2.1 percentage points for black and Hispanic boys, respectively. For all races/ethnicities, both higher male dropout rates and higher male high school grade repetition rates (which may culminate in dropout or high school completion) contribute substantially to the large gender gaps in the on-time high school graduation rate. These patterns are even more pronounced for black students.

The second column of each panel shows that the boy-girl disadvantage in high school completion is substantially smaller in more advantaged families. Conditional on race-ethnicity, the gender gap in on-time completions is 1.5 percentage points smaller for boys born to families with a one standard deviation higher SES level.[27] Accounting for family SES reduces the black excess boy-girl disadvantage in high school graduations from 6.7 to 5.0 percentage points (25 percent), and reduces the Hispanic excess boy-girl disadvantage from 2.1 to 1.3 percentage points (38 percent).

[26] See online Appendix Table O4 for results with individual covariates.

[27] These results are consistent with the evidence in Chetty et al. (2016b), who find a differential advantage of growing up in a higher income family for boys' relative to girls' college attendance, employment, and earnings outcomes. In contrast, recent papers using data from Denmark (Brenøe and Lundberg 2018) and the United States (Lundberg 2017) find that the differential negative effect of family disadvantage on boys' adolescent outcomes does not persist into adulthood. Our data do not permit us to assess post-high school outcomes.

To what degree do the upstream effects of SES on the gender gap in elementary and middle school disciplinary and academic outcomes plausibly account for the SES gradient in the on-time high school graduation gender gap? We are unable to fully answer this question since we cannot exhaustively account or control for all causal pathways. We can, however, offer a simple benchmark. We first estimate the predictive relationships between behavioral and academic outcomes that are observed in kindergarten and grades five through eight and on-time high school graduation, which is observed in grade 12.[28] We then apply our estimates of the effect of SES on these early behavioral and academic outcomes to the point estimates from this predictive model to obtain an implied effect of family disadvantage on high school graduations operating through these channels. We stress that these models for the predictive relationship between K–8 outcomes and subsequent high school graduations should be understood as descriptive; it is unlikely that the causal effect of family disadvantage on high school outcomes runs exclusively through K–8 behavioral and academic outcomes.

Table 6 reports this exercise. The first panel presents a linear probability regression of on-time high school graduation (H_i) on our standard set of covariates (child sex; race-ethnicity interacted with child sex; birth order; birth year; birth month; and the SES index based on mother's education, age and marital status at childbirth), augmented with a vector $\mathbf{T}_i$ of elementary and middle school behavioral and educational outcomes:

$$(8)\quad H_i = \alpha + \mathbf{T}_i'\pi + \beta_1 Boy_i + \beta_2(Boy_i \times Black_i) + \beta_3(Boy_i \times Hispanic_i) + \mathbf{X}_{1,i}'\lambda_1 + \mathbf{X}_{2,j(i)}'\lambda_2 + e_i.$$

To facilitate comparison across coefficients, we standardize each of the predictive variables in $\mathbf{T}$ to have mean zero and unit variance.

Estimates of equation (8) in Table 6 find a highly significant predictive relationship between early behavioral and educational outcomes and subsequent on-time high school completions. Whether entered into the regression individually (columns 1 through 5) or as a group (column 6), the data show that children who have higher absence and suspension rates and lower reading/math scores *during grades five through eight*, as well as lower rates of kindergarten readiness, are significantly less likely to complete high school within four years. Notably, the standardized effect sizes of early behavioral measures (absences and suspensions) are larger than the effect sizes of math and reading scores, which are, in turn, substantially larger than the effect sizes of kindergarten readiness.

To benchmark the economic magnitude of these coefficients relative to those above, we scale them by the estimated impact of SES on each behavioral and educational outcome and report them (in standardized form) in panel B of Table 6. The

[28] Our sample selection for this exercise is dictated by data availability: absences and suspensions are observed starting in the 2002–2003 school year; high school graduations are observed for the 1992 and 1993 birth cohorts. We use fifth through eighth grade outcomes for the 1992/1993 cohorts due to data limitations.

Table 6—The Relationship between Elementary and Middle School Behavioral and Academic Measures and On-Time High School Completion

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS Estimates: Kindergarten readiness, behavioral and academic measures, and on-time high school graduation						
Absence rate	−13.93					−10.41
	(0.16)					(0.17)
Suspension rate		−12.99				−8.30
		(0.18)				(0.19)
Math score			12.09			6.07
			(0.18)			(0.27)
Reading score				10.55		2.12
				(0.17)		(0.26)
Kindergarten readiness					3.11	0.42
					(0.17)	(0.16)
Observations				82,533		
Panel B. 100 × Standardized coefficients on Boy × SES from primary models						
Absence rate	−1.98					−1.98
Suspension rate		−6.01				−6.01
Math score			0.87			0.87
Reading score				0.31		0.31
Kindergarten readiness					3.34	3.34
Panel C. Implied contribution of cognitive/behavioral gender gaps in SES to high school graduation gender gaps						
	Absence rate	Suspension rate	Math score	Reading score	Kindergarten readiness	All
One SES σ	0.31	0.75	0.08	0.01	0.02	1.17
Black-white	0.30	0.72	0.08	0.02	0.02	1.12
Hispanic-white	0.13	0.32	0.03	0.01	0.01	0.50

Notes: Panel A of this table reports the results from regression of on-time high school graduation on absence rates, suspension rates, math scores, reading scores, and kindergarten readiness, respectively, each standardized with mean zero and unit variance. All regressions also include controls for gender, race, ethnicity, interactions between racial-ethnic categories and gender, the SES index (based on marital status at birth, maternal education, and maternal age at birth), child year and month of birth, and number of births. Panel B reports the coefficients on the interaction term Boy × SES in a standardized form based on the models estimated in Table 4 panels A and B, and Appendix Table A3. Due to data limitations necessitating the use of grades five to eight in panel A, we re-estimate our models from Table 4 panels A and B, using only grades five through eight for the purpose of this exercise. Panel C reports the implied contribution of the estimated SES gradient in the gender gap on the gender gap in high school graduation, by multiplying the coefficients from panels A and B, and scaling the result by the standard deviation in SES, black-white SES gap, and Hispanic-white SES gap.

final panel of Table 6 reports the implied impact of SES on high school graduations scaled by three metrics: a one standard deviation increment to SES (row 1); the mean white-black SES differential (row 2); and the mean white-Hispanic SES differential (row 3).

There is a large implied impact of family disadvantage on the gender gap in high school graduations operating through early behavioral and educational outcomes. A one standard deviation reduction in family SES ($\sigma = 1.51$) is predicted to raise the boy-girl high school graduation deficit by approximately 1.17 percentage points. Since cross-group differences in SES are substantial, this mechanism also contributes to observed race and ethnic gaps in on-time high school graduations. The mean black-white and Hispanic-white SES differentials of -1.44σ and -0.64σ, respectively, can explain 1.12 percentage points of the 6.7 percentage point excess boy-girl deficit in HS graduations among blacks and 0.50 percentage points of the 2.1 percentage point excess boy-girl deficit among Hispanics.

This estimated contribution operating through upstream channels is more than half as large as the direct effect of SES on the excess black-white gender gap in high school completions that we estimate in Table 5. The primary channel through which this effect appears to operate is behavioral outcomes. Variation in the gender gaps in reading and math achievement and in kindergarten readiness make a negligible contribution to the gender gap in HS graduations, jointly accounting for only 9.4 percent of the total explained by the fifth through eighth grade measures (versus 90.6 percent explained by absences and suspensions). A plausible interpretation of this finding is that, at least for boys, it is a behavioral rather than formal skills deficits that inhibit high school completion, as posited by Heckman and Kautz (2012). Indeed, of the five measures considered, suspensions play the largest role in explaining the boy-girl deficit in high school completions (panel C). Similarly, when we estimate models akin to panel A of Table 6 for high school dropout and 5+ years of high school attendance, we find that three-quarters of the estimated impact of behavioral measures on on-time graduation operates through differences in the likelihood of dropout.

C. *Oaxaca Restrictions*

Our econometric framework implicitly imposes the restriction that the differential impact of SES on boys' and girls' behavioral and educational outcomes is comparable across race-ethnicity groups, as noted in Section IIA. If the impact of SES on the gender gap differed substantially by demographic group, this would complicate interpretation of an SES "effect" on the cross race-ethnicity gender gap.

We test this restriction by estimating an augmented version of equation (7) that includes all second and third-level interactions between gender, SES, and race-ethnicity:

$$(9)\quad Y_i = \alpha + \beta_1 Boy_i + \beta_2(Boy_i \times D_i) + \beta_3(Boy_i \times Race_i) + \beta_4(Race_i \times D_i) + \beta_5(Boy_i \times Race_i \times D_i) + \mathbf{X}'_{1,i}\lambda_1 + \mathbf{X}'_{2,j(i)}\lambda_2 + e_i,$$

where D_i is the SES (disadvantage) measure, and we suppress notation of main effects of race-ethnicity and disadvantage. The assumption that SES affects minority and nonminority boys equally implies that $\beta_5 \simeq 0$ in the equation above. Appendix Tables A4 and A5 present these augmented specifications for the medium-run and high school graduation outcomes, respectively.

Focusing first on school absences in Appendix Table A4, we find no evidence of cross-race-ethnicity variation in the SES gradient in the gender gap. A Wald test for cross-race-ethnicity heterogeneity in the SES gradient in the gender gap fails to reject the null at $p = 0.29$. For reading and math scores, we find a modestly positive SES gradient in the gender gap for white and Hispanic children—reading and math outcomes for boys rise differentially with SES—but this gradient is modestly negative among black children. Finally, for suspensions, there is quantitatively and statistically significant heterogeneity in slopes. The negative coefficient on Boy $\times$ SES in column 2 of panel B implies that a one standard deviation increase in SES reduces the suspension rate of white boys relative to girls by 2.48 percentage points. The positive third-level interaction of 0.97 among boy, black, and SES in this model indicates that SES has a smaller beneficial effect $(-1.51 = -2.48 + 0.97)$ on the gender suspension gap among black children than white children. We find a similar pattern for Hispanics. While these cross-group discrepancies in slopes do not change the qualitative pattern of findings, they nevertheless suggest that part of the protective effect of SES in reducing suspensions among white males is not present for minority males. A plausible interpretation of this pattern, though not one that we are able to explore in our data, is that public school suspension policies are differentially punitive toward non-white males.

Using the estimates from equation (9), we can reconsider the contribution of family SES to the excess gender gap among black children relative to white children. Our previous estimates that constrain the Boy $\times$ SES coefficient to be constant across race/ethnic groups (from Table 4A) imply that switching from the SES level of the average black family to the SES level of the average white family (an increase of 1.44σ) would close the black-white excess gender gap in suspension rates by 3.12 percentage points (69 percent). When we allow the SES gradient in the gender gap to differ across race/ethnic groups, the average SES difference between black and white children implies a reduction in the excess gap of 2.17 percentage points. Although this calculation suggests that a smaller benefit of family advantage is imparted to minority relative to white boys, average black-white SES differences can still explain nearly 50 percent of the excess gender gap in suspensions among black children.

Appendix Table A5 applies our test of the Oaxaca restrictions to on-time high school graduation, high school grade repetition, and high school dropout. Because we do not have all components of the SES variable for the high school age cohorts in our sample, we use the alternative SES index that omits the indicator for Medicaid funded births and include third-level interactions between SES, race/ethnicity, and boy.[29] For on-time high school graduation, we cannot reject the null of no heterogeneity in the family SES gradient in the gender gap ($p \geq 0.97$). Among those not completing high school on time, we find some evidence that Hispanic boys from higher SES families are differentially likely to repeat a grade rather than drop out of high school.[30]

In light of the heterogeneity among race and ethnic groups in the relationship between family SES and the gender gap in academic and behavioral outcomes, we also revisit the contribution of SES to the gender gap in high school completion through these medium-term measures. For this calculation, we additionally modify equation (8)—the association between academic and behavioral outcomes and high school completion—to include interactions with race and ethnic categories. The results of these calculations are reported in online Appendix Table O5. The effect of family

[29] To confirm that this SES index is valid, we have compared the alternative SES measure (using fewer variables) with the primary (all variables) SES measure above within the subsample in which both measures can be calculated. Reassuringly, we obtain very similar estimated SES main effects and Boy $\times$ SES interactions (and slightly larger standard errors) in models for absences, suspensions, and math and reading scores.

[30] Murnane (2013) observes that black and Hispanic students are more likely than whites to complete high school in five-plus years. Thus, race gaps in eventual high school completion are smaller than race gaps in on-time graduation.

SES, as mediated through medium-term outcomes, explains less of the excess gender gap in high school completion among black relative to white students (0.44 instead of 1.12 percentage points from our primary models). While the *direct* contribution of family SES to the gender gap in high school graduation is comparable across race and ethnic groups (seen in Appendix Table A5), the *indirect* contribution of family SES through our academic and behavioral outcomes is substantially attenuated for black relative to white students. Of the various outcomes we consider, suspensions continue to play the largest role in mediating the effects of family SES. In interpreting these results, we place the highest weight on the *direct* effect of SES on high school graduations—a market-relevant outcome—and view early behavioral and academic outcome as mediators rather than end points. Given that we can observe high school graduations only for the oldest two cohorts in our sample, however, we maintain our primary empirical focus on these mediating measures.

IV. Assessing the Role of Latent Gaps

We assess whether there are latent differences between boys and girls that vary with family socioeconomic status that might give rise to the family SES gradient in the gender gap estimated above. To do so, we estimate the relationship between family disadvantage and the gender gap in the strongest available measure of neonatal health: birthweight. A large medical and economic literature summarized above finds birthweight to be robustly predictive of subsequent health, cognitive development, and labor market outcomes. In our sample, there are substantial birthweight differences across our three main demographic groups, as documented in Appendix Table A6. At birth, white children weigh an average of 250 grams (approximately 7.4 percent) more than black children, and about 90 grams more than Hispanic children. There are also substantial differences across these demographic groups in a number of other measures of birth outcomes.

Table 7 reports estimates of equations (6) and (7) for birthweight and an indicator for maternal health problems at birth to assess whether the gender gap in neonatal outcomes varies systematically with family disadvantage. If boys born to disadvantaged mothers are systematically less healthy than girls born to disadvantaged mothers, this could be reflected in birthweight. Alternatively, if maternal healthcare utilization or health are affected by (or correlated with) the gender of the fetus—perhaps due to son preference, as in Dahl and Moretti (2008)—our measure of maternal health might detect such a relationship. For each birth outcome, we report the same four specifications used in our main analysis. The first column reports conditional mean birthweight (or other birth outcome) differences between boys and girls, overall and by race-ethnicity, conditional on family disadvantage main effects. In the case of birthweight, the coefficient on the boy main effect indicates that white newborn boys weigh on average 121 grams more than white newborn girls. Despite the substantially lower average birthweight of black boys and girls, the black boy-girl birthweight gap is essentially identical to that of whites. We also find a modest mean difference in the gender-racial gap in birthweight among Hispanic births relative to white births of -16 grams.

The second column of each panel tests for a relationship between family disadvantage and the gender gap in birth outcomes by adding a set of interactions between

Table 7—Testing for a Family Disadvantage Effect on the Gender Gap in Neonatal and Maternal Health

	Panel A. Infant birth weight (g)				Panel B. Maternal health problems (percent)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Boy × Mother years of education		0.64				0.12		
		(0.71)				(0.06)		
Boy × Married		0.33				0.27		
		(3.27)				(0.26)		
Boy × Non-Medicaid birth		6.97				−0.10		
		(3.22)				(0.26)		
Boy × Mother age at birth		0.13				−0.03		
		(0.26)				(0.02)		
Boy × SES index			3.04	1.36			0.09	0.08
			(0.88)	(1.98)			(0.07)	(0.20)
Boy	120.60	104.11	119.83	124.16	0.16	−0.80	0.15	−0.09
	(1.49)	(8.47)	(1.52)	(3.44)	(0.12)	(0.68)	(0.12)	(0.35)
Boy × Black	−0.66	3.29	3.68	−3.55	−0.02	0.09	0.09	0.21
	(2.95)	(3.32)	(3.24)	(7.15)	(0.24)	(0.26)	(0.26)	(0.68)
Boy × Hispanic	−16.17	−14.28	−14.41	−10.78	0.03	0.06	0.11	−0.28
	(3.88)	(3.92)	(3.92)	(8.52)	(0.31)	(0.31)	(0.31)	(0.84)
Sibling fixed effects	No	No	No	Yes	No	No	No	Yes
Mean of Y		3,320		3,325		26.33		24.88
Number of children		796,701		301,298		796,701		301,298

Notes: This table reports estimates from regression models where the dependent variables are birthweight in grams and maternal health problems. Maternal health problems during pregnancy is equal to 100 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 4,000+ grams; previous preterm or small for gestational age infant; renal disease; RH sensitization; uterine bleeding; other specified health problem). All columns include controls for child year and month of birth, maternal race-ethnicity, and birth order within family. Columns 1 and 2 further control for maternal education, maternal age, non-Medicaid birth indicator, and marital status at birth. Columns 3 and 4 control for the main effect of the SES index, which is constructed as the first component of a principal components analysis of years of maternal education, maternal age, non-Medicaid birth indicator, and marital status at birth. Column 4 includes mother fixed effects. Robust standard errors are utilized in columns 1–3 and are clustered at mother level in column 4. See online Appendix Table O6 for an expanded set of results.

child sex and mother's age/education, the proxy for poverty, and mother's marital status at childbirth. The boy-girl difference in birthweight is trivially (and not statistically significantly) larger for births where the mother has more years of education, and is similarly unrelated to maternal age or marital status. We do observe a statistically significant male-favorable (i.e., positive) weight differential for non-Medicaid births, but this difference is only seven grams. The third column of each panel subsumes these four interaction terms into our single composite SES measure described above. We estimate a small but statistically significant relationship between family advantage and the gender gap in birthweight: a one standard deviation increase in the SES index (an increment of 1.51) predicts a three gram rise in the boy-girl differential in birthweight, a magnitude that is of negligible educational and health relevance.[31]

[31] Figlio et al. (2014) estimate that a 1,000g increase in birthweight is associated with a 0.19 standard deviation increase in test scores in grades three through eight. The three gram differential we estimate in Table 7 is 0.3 percent as large, implying a hypothetical impact on test scores on the order of three ten-thousandths of a standard deviation.

The results in Table 7 panel B show no evidence of a gender gap (or excess gender gaps) in maternal health problems. In addition, there is little evidence that the gender gap in maternal health varies systematically with measures of family disadvantage. In online Appendix Table O7 we report results for two additional neonatal outcomes: the logarithm of birthweight, and a composite health-at-birth index.[32] For both outcomes, we find no economically or robustly statistically significant relationship between family advantage and the gender gap in neonatal well-being.[33] Online Appendix Figure O1 summarizes these patterns by presenting scatter plots of the relationship between family SES and the gender gaps in neonatal health outcome. Across multiple outcome measures, the relationship between family advantage and the gender gap in infant health is of trivial economic magnitude and, in three of four cases, is statistically insignificant.

As a further test of unconfoundedness we report in online Appendix Table O8, a set of augmented estimates for the relationship between SES and the gender gap in behavioral and educational outcomes that control for the log of birthweight and its interaction with gender (in addition to all prior covariates). If there is any residual relationship between family disadvantage, infant birthweight, and the gender gap in subsequent behavioral or educational outcomes, these augmented estimates should account for that relationship, and our main estimates should be attenuated accordingly. Consistent with expectations, children with higher birthweight attain higher math and reading scores during third through eighth grades, though they also have higher suspension rates. Inclusion of infant birthweight has no discernible impact, however, on the estimated SES gradient in the gender gap in behavioral or educational outcomes.[34]

V. Exploring Mechanisms: Schools and Neighborhoods

Family disadvantage may amplify the female-favorable gap in childhood outcomes not exclusively because boys are differentially affected by family environment, but also because the neighborhoods and schools in which disadvantaged children are raised are particularly adverse for boys. Boys may be more vulnerable to the risks of violence and gang activity in low-SES neighborhoods and schools, or boys may be treated more harshly by authority figures in these settings, perhaps facing greater disciplinary and criminal sanctions from teachers and police. We formally test whether family disadvantage operates primarily through these channels rather than directly through family environment per se. While schools and neighborhoods appear less important for explaining the gender gap than the direct effect of family environment itself, the results in this section lend broad support to the hypothesis that early outcomes of boys—particularly, behavioral and disciplinary

[32] This index is constructed from a principal components analysis of birthweight, gestational age, one and five minutes Apgar scores (0–10 scale) as well as indicators for adequate prenatal care, maternal health problems in pregnancy, complications of labor and delivery, abnormal conditions at birth, and congenital anomalies.

[33] The logarithmic birthweight parameterization tests whether SES affects the *proportional* rather than the *level* difference in the gender gap in birthweight. This distinction is meaningful since boys weigh significantly more than girls across all demographic groups.

[34] We further estimated models that control for a wider variety of birth outcome measures and their interactions with infant gender, and these additions leave our results fundamentally unchanged.

Table 8A—Determinants of the Gender Gap in Behavioral Outcomes: Neighborhood Income, School Quality, and Economic Mobility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Absence rate (percent)							
Boy × SES index		−0.063	−0.058	−0.063	−0.049	−0.048	
		(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	
Boy × Income in US $10,000			−0.018			−0.007	−0.018
			(0.008)			(0.008)	(0.008)
Boy × Mobility				−0.022		−0.056	−0.053
				(0.050)		(0.051)	(0.051)
Boy × School quality					−0.004	−0.004	−0.005
					(0.001)	(0.001)	(0.001)
Boy	0.173	0.188	0.271	0.182	0.395	0.407	0.517
	(0.012)	(0.013)	(0.039)	(0.017)	(0.040)	(0.048)	(0.047)
Boy × Black	0.238	0.150	0.138	0.149	0.122	0.117	0.158
	(0.024)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.026)
Boy × Hispanic	0.064	0.025	0.022	0.024	0.028	0.025	0.050
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Mean of *Y*				5.083			
Number of children				754,399			
Panel B. Suspension rate (percent)							
Boy × SES index		−2.13	−1.99	−2.12	−1.81	−1.77	
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	
Boy × Income in US $10,000			−0.49			−0.20	−0.62
			(0.04)			(0.04)	(0.04)
Boy × Mobility				0.51		−0.34	−0.20
				(0.25)		(0.26)	(0.26)
Boy × School quality					−0.09	−0.09	−0.13
					(0.00)	(0.00)	(0.00)
Boy	7.98	8.48	10.80	8.60	13.51	14.10	18.13
	(0.05)	(0.06)	(0.18)	(0.08)	(0.19)	(0.23)	(0.22)
Boy × Black	4.51	1.52	1.19	1.53	0.84	0.73	2.23
	(0.14)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
Boy × Hispanic	−0.49	−1.83	−1.91	−1.80	−1.74	−1.80	−0.88
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
Mean of *Y*				12.75			
Number of children				754,399			

Note: See note to Table 8B.

outcomes—are differentially impacted by disadvantage, whether manifested in family disadvantage, neighborhood quality, or school quality.

For this analysis, we augment our data with measures of school and neighborhood quality, described in Section I. Table 8A presents estimates for absences and suspensions. Columns 1 and 2 replicate our baseline specification for the gender gap in school absences, while subsequent columns augment this specification with school and neighborhood controls, in each case interacted with gender.[35] Columns 3 and 4 show that the gender gap in absences is lower in higher income zip codes and areas with higher economic mobility, though the latter relationship is not statistically significant. The gender gap in absences is also smaller for children who attend higher quality Florida public schools relative to those attending lower quality schools (column 5).

[35] Samples (and hence point estimates) differ slightly from our main estimates in Table 4a due to the fact that school quality measures are missing for a small subset of schools. We report estimates only for the sample of singletons, but the conclusions hold with inclusion of sibling fixed effects. See Autor et al. (2016a) for the results.

Moving from the twenty-fifth to the seventy-fifth percentile in school quality is predicted to reduce the gender gap in absenteeism by 0.002 percentage points. The sixth column includes all three school and neighborhood measures simultaneously. School quality remains robustly predictive of the gender gap in absences and the economic mobility coefficient increases but remains statistically insignificant.

We use two metrics to interpret the magnitude of these relationships: how much of the SES gradient in that gap do they explain; and how much of the race-ethnicity gender gap in absences do they explain? The first question can be answered by comparing the coefficients on the Boy $\times$ SES interaction term across columns of Table 8A, while including versus excluding the neighborhood and school quality interaction terms. The coefficient on Boy $\times$ SES of -0.063 in the column 2 baseline specification falls in magnitude to -0.048 in the final specification that accounts for both school and neighborhood quality. Thus, 24 percent $(1 - 0.048/0.063)$ of the effect of SES on the gender gap in absences is accounted for by measures of school quality and neighborhood economic attributes. The remainder is implicitly accounted for by family-level effects that operate within schools and neighborhoods. Of course, this decomposition provides a lower bound on the explanatory power of schools and neighborhoods: more complete measures of school and neighborhood quality could potentially explain more. We address this issue in online Appendix Table O9, summarized below.

The answer to the second question—what share of the race-ethnicity gender gap is explained by family, school, and neighborhood—is summarized in Figure 3. As reported in Table 8A, the regression adjusted excess gender gap in school absences among blacks relative to whites is 0.24 percentage points. Controlling for the differential impact of family advantage on boys accounts for 37 percent of this gap. Adding neighborhood income explains an additional 0.01 percentage points (42 percent in total), while controlling for school quality and neighborhood economic mobility increases this explanatory power an additional 9 percentage points (to 51 percent). Logically, school and neighborhood controls have substantially greater incremental explanatory power for the excess gender gap in absences when SES is excluded from the model: controlling for these factors alone, absent the gender-SES interaction, explains 34 percent of the excess black gender gap in absences versus 37 percent for SES alone (column 7 of Table 8A). In summary, 73 percent of the explained component of the excess gender gaps $(1 - [0.14/(0.14 + 0.37)])$ appears to operate *within* schools and neighborhoods. The high correlation between the two sets of variables means that either set can proximately account for a substantial share of the explained excess gap. Nevertheless, inclusion of family SES in the neighborhoods/schools specification substantially improves the model's ability to account for the excess gap.

Panel B of Table 8A implements this exercise for grades three through eight suspensions. Here we find that both neighborhood income and economic mobility—as well as school quality—are significant predictors of the boy-girl gap in suspensions, with a smaller female-favorable gap in higher quality schools, and wealthier neighborhoods.[36] When accounting for all school and neighborhood measures

[36] Unexpectedly, we find, in column 4, a larger female-favorable gap in more economically mobile neighborhoods. But this relationship flips sign (in the anticipated direction) when we condition on other neighborhood and school variables.

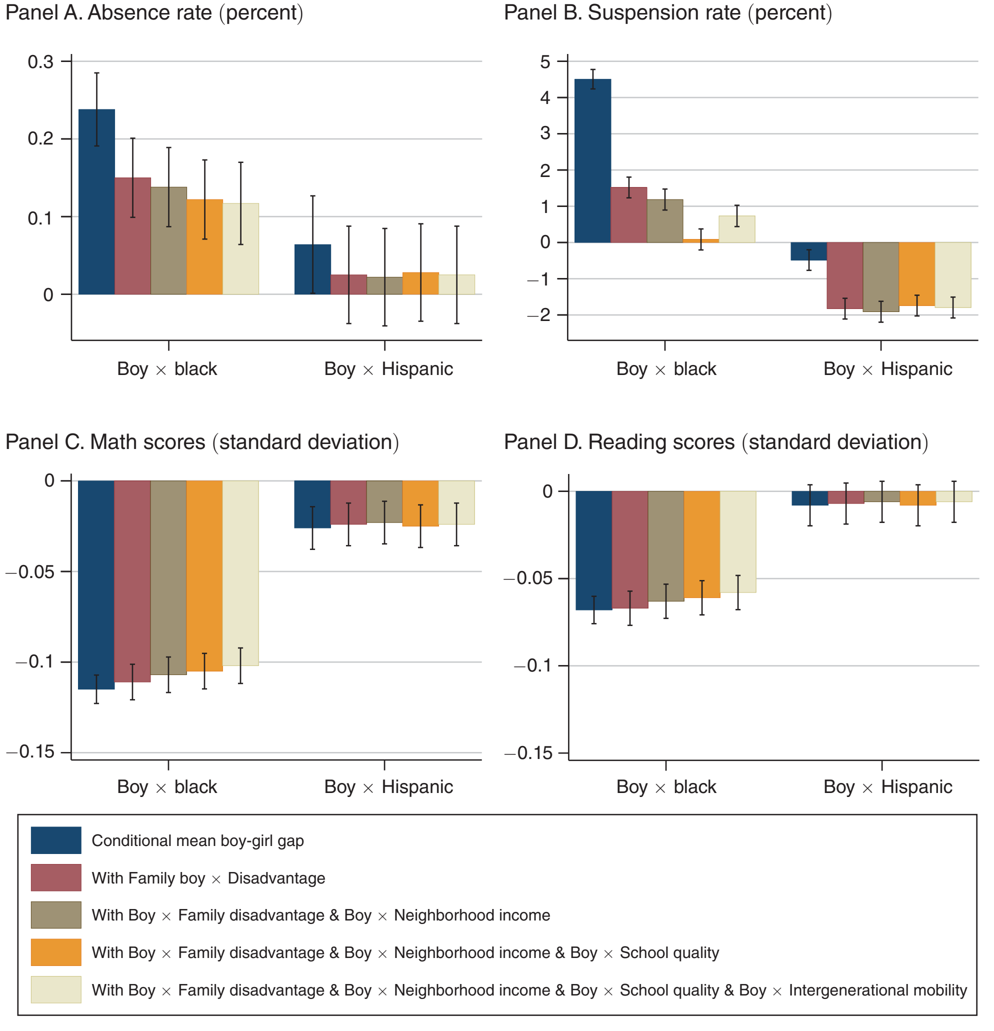

Figure 3. Gender Gaps in Behavioral and Academic Outcomes: Observed and Explained

Notes: This figure plots the regression-adjusted, race-ethnicity gender gap in absence rates, suspension rates, FCAT math and reading scores from regression models in Tables 8A and 8B. Ninety-five percent confidence intervals are represented by whiskers on each bar.

simultaneously, intergenerational mobility no longer significantly contributes to the gender gap in suspensions. Similar to the results for absences, the inclusion of all three neighborhood and school quality measures only minimally affects the family-level impact of SES on the gender gap: the SES × boy interaction term falls by less than one-fifth when conditioning on these detailed measures. As visually depicted in panel B of Figure 3, family disadvantage alone accounts for 66 percent of the excess gender gap among black relative to white children, with school and neighborhood quality accounting for an incremental 18 percentage points. The final column of Table 8A indicates that when the gender-SES interaction is excluded from the statistical model, schools and neighborhoods alone can explain 51 percent of the excess gender gap in suspensions among blacks relative to whites.

Table 8B—Determinants of the Gender Gap in Academic Outcomes: Neighborhood Income, School Quality, and Economic Mobility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Mathematics scores (standard deviation)							
Boy × SES index		0.003	0.001	0.003	−0.000	−0.001	
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
Boy × Income in US $10,000			0.007			0.004	0.004
			(0.002)			(0.002)	(0.002)
Boy × Mobility				0.002		0.012	0.012
				(0.009)		(0.009)	(0.009)
Boy × School quality					0.001	0.001	0.001
					(0.000)	(0.000)	(0.000)
Boy	0.041	0.041	0.010	0.041	−0.007	−0.019	−0.017
	(0.002)	(0.002)	(0.008)	(0.003)	(0.007)	(0.009)	(0.009)
Boy × Black	−0.115	−0.111	−0.107	−0.111	−0.105	−0.102	−0.102
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Boy × Hispanic	−0.026	−0.024	−0.023	−0.024	−0.025	−0.024	−0.023
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mean of *Y*				0.066			
Number of children				753,356			
Panel B. Reading scores (standard deviation)							
Boy × SES index		0.001	−0.000	0.001	−0.002	−0.002	
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
Boy × Income in US $10,000			0.005			0.004	0.003
			(0.001)			(0.002)	(0.002)
Boy × Mobility				0.022		0.031	0.031
				(0.009)		(0.009)	(0.009)
Boy × School quality					0.001	0.001	0.001
					(0.000)	(0.000)	(0.000)
Boy	−0.141	−0.142	−0.166	−0.136	−0.182	−0.189	−0.185
	(0.002)	(0.002)	(0.007)	(0.003)	(0.007)	(0.009)	(0.008)
Boy × Black	−0.068	−0.067	−0.063	−0.066	−0.061	−0.058	−0.057
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Boy × Hispanic	−0.008	−0.007	−0.006	−0.006	−0.008	−0.006	−0.004
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mean of *Y*				0.079			
Number of children				753,356			

Notes: These tables present results from regression models where the dependent variables are the absence rate, suspension rate, and standardized FCAT math and reading test scores, from grades three through eight. All columns include controls for child year and month of birth, maternal race-ethnicity, and birth order within family, the SES index, income of the zip code of residence at the time of birth, the Chetty-Hendren mobility measure, and the Florida school quality measure. Column 2 additionally includes the interaction of boy and the SES index. Column 3 further includes the interaction of boy and income of the zip code of residence at the time of birth. Column 4 includes the interaction of boy and the Chetty-Hendren mobility measure. Column 5 includes the interaction of boy and school quality. Column 6 includes the interactions of all four measures with boy simultaneously. Column 7 replicates column 6 but excludes interaction between boy and SES index. The sample excludes families for whom zip code, Chetty-Hendren mobility measure, or school quality are not known. Standard errors are clustered at the child level.

Table 8B presents the corresponding analysis for math and reading scores. Relative to the behavioral outcomes, these academic indicators are less central to our analysis since they have comparatively modest relationships to high school completion, grade repetition, and dropout. They also exhibit a more modest SES gender gradient, as shown above. Qualitatively, we find a similar pattern: greater neighborhood income and higher school quality significantly reduce the boy-girl disadvantage in early academic performance, while we do not find a consistently significant effect of neighborhood economic mobility. Accounting for neighborhood and school quality attenuates the relationship between SES and the gender gap in math and reading and in most specifications the coefficient is statistically insignificant. However, *none* of these measures (individually or in combination) accounts for even a fifth of the excess gender gap in academic outcomes among black relative to white children.

As a final robustness test, in online Appendix Table O9, we estimate a set of companion models for absences, suspensions, and math and reading scores in which we non-parametrically account—to the extent of the variation available in our dataset—for the role of neighborhoods and schools by including a full set of fixed effects for zip codes and schools interacted with boy. Relative to the Table 8A models that control directly for neighborhood and school quality, we find that adding nonparametric school and neighborhood controls has almost no discernible effect on the estimated impact of family disadvantage on the gender gap in absences and suspensions. For reading and math scores, the attenuation is of a similar magnitude to the parametric approach, and the interaction of gender and SES is in most specifications no longer significant. Family disadvantage appears, however, to account for only a modest component of the gender gap in test scores, as noted above, and test scores in turn play only a modest role in the gender gap in high school graduations. Although we do not have information on neighborhoods for the 1992 and 1993 cohorts, we also investigate the role of school quality for on-time high school completion, completing high school in five-plus years, and dropping out of high school; whether school quality differentially affects boys' outcomes; and whether these differential effects explain our prior estimates of the effect of family disadvantage on the gender gap. These results, reported in Appendix Table A7, show that higher quality schools differentially raise on-time graduations and reduce dropout among boys relative to girls. Accounting for this school quality channel does not attenuate the coefficient on the Boy $\times$ SES interaction and explains a small fraction of the excess gender gap in high school outcomes among blacks and Hispanics relative to whites.

VI. Conclusions

This paper investigates whether family disadvantage exerts a differential effect on the developmental and educational outcomes of boys relative to girls. Utilizing a unique dataset of all Florida births between 1992 and 2002 linked to public school records, we find that family disadvantage disproportionately negatively affects the behavioral and academic outcomes of school-age boys relative to girls. The differential effect of family disadvantage on the outcomes of boys relative to girls is already evident by the time of kindergarten entry, is further manifested in behavioral and educational gaps in elementary and middle school performance, and crystalizes into sharp differences in high school graduations by age 18. These SES gradients are especially strong for behavioral outcomes measured during the third through eighth grades. Our results imply that a sizable portion of the documented minority-white difference in educational and behavioral gender gaps can be attributed to higher degrees of family disadvantage among minority families. Most of this effect appears to operate through behavioral development, seen in our data in the differentially high rate of absences and suspensions among low-SES boys relative to low-SES girls. While family disadvantage also appears to inhibit boys' early math and reading achievement, the effect sizes are much smaller *and* these academic outcomes are far less predictive of high school non-completion than are behavioral outcomes observed at the same ages.

The effect of family disadvantage on the gender gap may accrue through two primary channels. One is that the skills development of boys and girls respond differently to the same stimuli. An alternative, non-mutually exclusive, explanation is that parental investments in boys versus girls differ systematically according to family disadvantage. For example, parents in low-SES households, which are disproportionately female-headed, may spend relatively more time mentoring and interacting with daughters than sons (Lundberg, Pabilonia, and Ward-Batts 2007; Baker and Milligan 2013; Bertrand and Pan 2013; Gayle, Golan, and Soytas 2015), or similarly, parents in high-SES households may make larger compensatory investments in sons than daughters.[37] Our data do not allow us to evaluate the relative importance of these causal channels, though the evidence in Bertrand and Pan (2013) does not provide strong support for the differential investment hypothesis. We view this as a worthy topic for further research.

While our paper is agnostic about the specific channels through which postnatal exposure to family disadvantage generates the observed gender gap in outcomes, we are able to test and, for the most part, reject two natural competing explanations for our main findings. One explanation for the SES gradient in the gender gap is a "fetal origins" hypothesis, in which the SES gradient in potential outcomes is imparted prior to birth. Across a number of measures of neonatal health, we detect little evidence supporting this hypothesis. The second explanation for the SES gradient in the gender gap is an environmental factors hypothesis. We document that neighborhood quality makes a small contribution to the SES gradient in the gender gap, while school quality matters more. Nevertheless, accounting non-parametrically for the differential impact of schools and neighborhoods on boys relative to girls shrinks the estimated impact of family disadvantage on the sibling gender gap in behavioral outcomes by less than a quarter. These results suggest that the divergent behavioral and cognitive development of boys relative to girls in low-SES versus high-SES families reflects the effect of the postnatal family environment on child development and this effect is largely distinct from (or in addition to) the impact of other environmental factors such as schools and neighborhoods. Though not observable in our data, we suspect that these early differences in behavioral and educational outcomes continue into adulthood, as boys and girls exit the compulsory school system and matriculate into employment, higher education, and potentially parenthood.

[37] We have examined this hypothesis using parental time use data from the American Time Use Survey as well as data from the Florida State Department of Education on whether parents send children to preschool, an indicator of early childhood investment. With regard to parental time use, we find little support for differential investment in boys versus girls that varies systematically with SES. For the likelihood of sending a child to prekindergarten programs, we find that boys are more likely to attend prekindergarten programs, and particularly so in relatively advantaged families.

Appendix Tables

Table A1—Descriptive Statistics: Behavioral and Academic Outcomes in Kindergarten and Grades 3–8

	All (1)	White non-Hispanic (2)	Black non-Hispanic (3)	Hispanic (4)
Panel A. Means				
Kindergarten ready (percent)	86.19	88.94	80.49	82.88
	(34.50)	(31.37)	(39.63)	(37.67)
	361,644	233,168	89,864	38,612
Absence rate (percent)	5.11	5.12	5.10	5.12
	(5.47)	(5.34)	(5.89)	(5.20)
	3,415,396	2,196,415	865,577	353,404
Suspension rate (percent)	12.82	8.87	24.01	9.93
	(33.43)	(28.43)	(42.71)	(29.90)
	3,415,396	2,196,415	865,577	353,404
Reading score (standard deviation)	0.08	0.29	−0.42	0.00
	(0.96)	(0.92)	(0.89)	(0.92)
	3,734,111	2,392,519	947,048	394,544
Math score (standard deviation)	0.06	0.27	−0.45	−0.01
	(0.96)	(0.90)	(0.93)	(0.92)
	3,724,571	2,385,927	945,707	392,937
Panel B. Boy-girl differences				
Kindergarten ready (percent)	−6.07	−5.65	−7.75	−5.91
	(0.11)	(0.13)	(0.26)	(0.38)
	361,644	233,168	89,864	38,612
Absence rate (percent)	0.26	0.17	0.47	0.25
	(0.01)	(0.01)	(0.01)	(0.02)
	3,415,396	2,196,415	865,577	353,404
Suspension rate (percent)	8.93	8.01	12.64	7.60
	(0.04)	(0.04)	(0.09)	(0.10)
	3,415,396	2,196,415	865,577	353,404
Reading score (standard deviation)	−0.15	−0.14	−0.22	−0.15
	(0.00)	(0.00)	(0.00)	(0.00)
	3,734,111	2,392,519	947,048	394,544
Math score (standard deviation)	0.02	0.04	−0.08	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
	3,724,571	2,385,927	945,707	392,937

Notes: This table reports kindergarten readiness rates (multiplied by 100), mean absence rates (multiplied by 100), ever suspended rates (multiplied by 100), and standardized FCAT reading and math scores, by race-ethnicity category, from the sample of non-twin singletons. Absence rates, suspension rates, and test scores are for birth cohorts 1994–2002 and span grades three through eight, with each child contributing up to one observation per grade observed in each year. Kindergarten readiness rates are computed from cohorts 1994–1996 and 2000–2002. In panel A, standard deviations are reported in parentheses; in panel B, standard errors are reported in parentheses. Numbers of observations are recorded beneath each standard deviation/error.

Table A2—High School Completion Rates (1992 and 1993 Birth Cohorts)

	All	White non-Hispanic	Black non-Hispanic	Hispanic
	(1)	(2)	(3)	(4)
Panel A. Means (percent)				
On-time graduate	70.42	72.70	64.63	68.00
5+ years	12.75	10.39	19.15	13.78
Dropout	16.83	16.91	16.22	18.22
Panel B. Boy-girl differences				
On-time graduate	−7.43	−5.69	−12.74	−8.04
	(0.23)	(0.27)	(0.48)	(0.90)
5+ years	4.40	3.78	6.83	4.20
	(0.17)	(0.18)	(0.39)	(0.67)
Dropout	3.02	1.91	5.91	3.85
	(0.19)	(0.22)	(0.37)	(0.75)
Number of students	161,537	111,454	39,517	10,566

Notes: This table reports on-time high school graduation rates, continuation in high school rates, and high school dropout rates, by race-ethnicity category, from the sample of non-twin singletons who were observed in Florida public schools until at least ninth grade. All rates and boy-girl differences are multiplied by 100. On-time high school completion indicates a high school diploma within 4 years of entering. 5+ years of high school indicates that a student is enrolled in high school more than 4 years after entry but had not dropped out. High school dropout indicates that an individual does not earn a high school diploma and is no longer enrolled in high school 5+ years after entry. Standard errors for male-female contrasts are reported in parentheses in panel B.

Table A3—Family Disadvantage and the Gender Gap in Kindergarten Readiness

	Panel A. Singletons OLS		*Panel B. Siblings OLS*		*Panel C. Mother FE*	
	(1)	(2)	(1)	(2)	(1)	(2)
Boy × SES index		1.15		1.08		1.13
		(0.08)		(0.22)		(0.41)
Boy	−5.80	−6.03	−5.97	−6.09	−5.88	−5.99
	(0.12)	(0.13)	(0.37)	(0.38)	(0.70)	(0.71)
Boy × Black	−2.04	−0.46	−2.45	−0.86	−2.74	−1.09
	(0.28)	(0.30)	(0.70)	(0.80)	(1.31)	(1.49)
Boy × Hispanic	−0.36	0.33	−2.25	−1.52	−1.87	−1.10
	(0.38)	(0.38)	(1.02)	(1.04)	(1.93)	(1.96)
Mean of Y	85.82		83.05		83.05	
Number of children	396,074		55,419		55,419	

Notes: This table reports the results of regression models with and without sibling fixed effects where the dependent variable is kindergarten readiness, which takes on a value of 100 if the child is determined to be ready for kindergarten, and 0 otherwise. All columns include controls for child year and month of birth, maternal race-ethnicity, birth order within family or sibling composition, and the main effect of the SES index. Panel A uses the sample of non-twin singletons born 1994 to 1996 and 2000 to 2002. Panels B and C restrict the sample to matched siblings—children in families with two or more births from 1994 to 1996 or from 2000 to 2002. In panel A, we use robust standard errors. In panels B and C, standard errors are clustered at the mother level.

Table A4—Testing for Heterogeneous Effects of Family Disadvantage on the Gender Gap in Behavioral and Academic Outcomes among Race/Ethnic Groups

	Panel A. Absence rate (percent)		*Panel B. Suspension rate (percent)*		*Panel C. Math scores (standard deviation)*		*Panel D. Reading scores (standard deviation)*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Boy × SES × Black (A1)		0.02		0.97		−0.009		−0.009
		(0.02)		(0.10)		(0.003)		(0.003)
Boy × SES × Hispanic (A2)		0.02		0.73		−0.000		0.002
		(0.02)		(0.10)		(0.004)		(0.004)
Boy × SES index	−0.07	−0.08	−2.17	−2.48	0.005	0.006	0.002	0.004
	(0.01)	(0.01)	(0.03)	(0.04)	(0.001)	(0.002)	(0.001)	(0.002)
Boy × Black	0.17	0.18	1.48	2.18	−0.112	−0.120	−0.067	−0.075
	(0.03)	(0.03)	(0.14)	(0.15)	(0.005)	(0.005)	(0.004)	(0.005)
Boy × Hispanic	0.03	0.03	−1.80	−1.72	−0.023	−0.021	−0.007	−0.005
	(0.03)	(0.03)	(0.14)	(0.13)	(0.006)	(0.006)	(0.006)	(0.006)
Boy	0.21	0.21	8.58	8.65	0.033	0.033	−0.148	−0.148
	(0.01)	(0.01)	(0.05)	(0.06)	(0.002)	(0.002)	(0.002)	(0.002)
H_0 (A1 = A2 = 0)		0.29		0.00		0.02		0.01
Number of children	792,729		792,729		785,664		785,673	

Notes: This table presents results from regression models in which the dependent variables are absence rates, suspension rates, and FCAT math and reading scores. All columns include controls for child year and month of birth, maternal race-ethnicity, birth order within family, second-level interactions between race-ethnic groups and the SES index, and the main effect of the SES index. Standard errors are clustered at the child level. *p*-values associated with the Wald tests of coefficient equality are reported at the bottom of the table.

Table A5—Testing for Heterogeneous Effects of Family Disadvantage on the Gender Gap in High School Outcomes among Race/Ethnic Groups

	Panel A. On-time high school graduation (percent)	*Panel B. 5+ Years high school (percent)*	*Panel C. High school dropout (percent)*
Boy × Black × SES index	−0.11	0.55	−0.44
	(0.43)	(0.35)	(0.35)
Boy × Hispanic × SES index	0.00	1.33	−1.33
	(0.71)	(0.56)	(0.62)
Boy × SES index	1.47	−1.07	−0.40
	(0.22)	(0.16)	(0.20)
Boy × Black	−5.04	2.21	2.83
	(0.60)	(0.48)	(0.46)
Boy × Hispanic	−1.26	0.11	1.14
	(0.90)	(0.68)	(0.74)
Boy	−6.47	4.21	2.25
	(0.28)	(0.20)	(0.24)
H_0 (A1 = A2 = 0)	0.97	0.03	0.07
Number of children		161,537	

Notes: This table presents the results of regression models for the sample singletons born in 1992–1993. On-time high school completion takes on a value of 100 if a student obtains a high school diploma within 4 years of entering, and is 0 otherwise. 5+ years of high school takes on a value of 100 if the student is enrolled in high school more than 4 years after entry but has not yet dropped out, and is 0 otherwise. High school dropout takes on a value of 100 if a student does not earn a high school diploma and is no longer enrolled in high school 5+ years after entry, and is 0 otherwise. The SES index is constructed as the first component of a principal components analysis of years of maternal education, maternal age, and marital status at birth. All columns include main effects for race-ethnicity of mother, child year and month of birth, birth order, second-level interactions between race-ethnic groups and the SES index, and the main effect of the SES index. *p*-values associated with the Wald tests of coefficient equality are reported at the bottom of the table. Robust standard errors are in parentheses.

Table A6—Descriptive Statistics: Neonatal and Medium-Term Health Outcomes

	All	White non-Hispanic	Black non-Hispanic	Hispanic
	(1)	(2)	(3)	(4)
Panel A. Means				
Infant birth weight (g)	3,320	3,392	3,142	3,305
log birth weight	8.09	8.11	8.03	8.09
Maternal health issues (percent)	26.33	25.84	29.29	22.38
Health index (standard deviation)	−0.06	0.04	−0.35	0.01
Panel B. Boy-Girl Differences				
Infant birth weight (g)	122	121	120	105
	(1)	(2)	(3)	(4)
Log birth weight	0.04	0.04	0.04	0.03
	(0.00)	(0.00)	(0.00)	(0.00)
Maternal health issues (percent)	0.12	0.16	0.16	0.19
	(0.10)	(0.12)	(0.20)	(0.28)
Health index (standard deviation)	0.04	0.02	0.08	0.01
	(0.00)	(0.00)	(0.01)	(0.01)

Notes: This table presents summary statistics for children's at-birth health outcomes. Maternal health issues during pregnancy are equal to 100 if the mother suffered from any of a large set of chronic or pregnancy-related disorders during pregnancy or delivery: anemia; cardiac disease; acute or chronic lung disease; diabetes; genital herpes, hydramnios/oligohydramnios; hemoglobinopathy; chronic hypertension; pregnancy associated hypertension; eclampsia; incompetent cervix; previous infant 4,000+ grams; previous preterm or small for gestational age infant; renal disease; RH sensitization; uterine bleeding; other specified health problem. The prenatal health index is a first component from a PCA analysis using birth weight (grams), gestational age (weeks), one and five minutes Apgar scores (0–10 scale) as well as indicators for adequate prenatal care, maternal health problems in pregnancy, complications of labor and delivery, abnormal conditions at birth, and congenital anomalies. Numbers of observations for infant birthweight: 796,701 for all; 509,372 for white non-Hispanic; 200,812 for black non-Hispanic; and 86,517 for Hispanic. Sample sizes vary slightly (by less that 0.1 percent) for the other birth outcomes.

Table—A7 Determinants of the Gender Gap in High School Graduation: School Quality

	(1)	(2)	(3)	(4)
Panel A. On-time high school graduation				
Boy × SES index	1.35	1.29	1.12	1.27
	(0.20)	(0.21)	(0.21)	(0.22)
Boy × School quality		0.22		
		(0.01)		
Boy × Black	−4.82	−4.64	−3.19	−3.97
	(0.62)	(0.63)	(0.71)	(0.74)
Boy × Hispanic	−0.94	−0.90	0.18	−0.11
	(0.99)	(0.98)	(1.05)	(1.07)
Mean of *Y*	70.87			
Panel B. 5+ years of high school				
Boy × SES index	−0.83	−0.74	−0.71	−0.74
	(0.15)	(0.16)	(0.17)	(0.17)
Boy × School quality		−0.08		
		(0.01)		
Boy × Black	1.72	1.47	1.12	0.91
	(0.50)	(0.51)	(0.57)	(0.60)
Boy × Hispanic	−0.47	−0.48	−1.28	−1.03
	(0.75)	(0.75)	(0.82)	(0.83)
Mean of *Y*	13.00			
Panel C. High school dropout				
Boy × SES index	−0.51	−0.55	−0.41	−0.53
	(0.17)	(0.18)	(0.18)	(0.18)
Boy × School quality		−0.14		
		(0.01)		
Boy × Black	3.10	3.17	2.07	3.07
	(0.50)	(0.52)	(0.58)	(0.61)
Boy × Hispanic	1.41	1.38	1.09	1.14
	(0.82)	(0.82)	(0.88)	(0.89)
Mean of *Y*	16.14			
Number of children	137,806			
Grade eight school fixed effects	No	No	Yes	No
Modal school fixed effects	No	No	No	Yes

Notes: This table reports the results of regression models where the dependent variable is, in turn, on-time graduation, continuation in high school, and dropout, from the sample of non-twin singletons who were born in 1992 or 1993 and observed in Florida public schools until at least ninth grade. On-time high school completion takes on a value of 100 if a student obtains a high school diploma within 4 years of entering, and is 0 otherwise. 5+ years of high school takes on a value of 100 if the student is enrolled in high school more than 4 years after entry but has not yet dropped out, and is 0 otherwise. High school dropout takes on a value of 100 if a student does not earn a high school diploma and is no longer enrolled in high school 5+ years after entry, and is 0 otherwise. The SES index is constructed as the first component of a principal components analysis of years of maternal education, maternal age, and marital status at birth. Column 1 does not include any school controls, column 2 controls for average school quality experienced between grades three and eight and its interaction with gender, column 3 includes Grade 8 school fixed effects and their interaction with gender, and column 4 controls for modal school fixed effects (most commonly attended school between grades three and eight) and their interaction with gender. Robust standard errors are in parentheses.

REFERENCES

Aizer, Anna, and Janet Currie. 2014. "The Intergenerational Transmission of Inequality: Maternal Disadvantage and Health at Birth." *Science* 344 (6186): 856–61.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–72.

Almond, Douglas, and Lena Edlund. 2007. "Trivers–Willard at Birth and One Year: Evidence from US Natality Data 1983–2001." *Proceedings of the Royal Society B: Biological Sciences* 274 (1624): 2491–96.

Almond, Douglas, and Bhashkar Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy." *American Economic Journal: Applied Economics* 3 (4): 56–85.

Almond, Douglas, Kenneth Y. Chay, and David S. Lee. 2005. "The Costs of Low Birth Weight." *Quarterly Journal of Economics* 120 (3): 1031–83.

Almond, Douglas, Bhashkar Mazumder, and Reyn van Ewijk. 2015. "*In Utero* Ramadan Exposure and Children's Academic Performance." *Economic Journal* 125 (589): 1501–33.

Arcidiacono, Peter, Andrew Beauchamp, Marie Hull, and Seth Sanders. 2015. "Exploring the Racial Divide in Education and the Labor Market through Evidence from Interracial Families." *Journal of Human Capital* 9 (2): 198–238.

Aucejo, Esteban M., and Jonathan James. 2017. "Catching Up to Girls: Understanding the Gender Imbalance in Educational Attainment within Race." https://www.cob.calpoly.edu/undergrad/wp-content/uploads/sites/3/2017/06/paper1701.pdf.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes: Dataset." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20170571.

Autor, David H. 2014. "Skills, Education, and the Rise of Earnings Inequality among the 'Other 99 Percent.'" *Science* 344 (6186): 843–51.

Autor, David, and Melanie Wasserman. 2013. "Wayward Sons: The Emerging Gender Gap in Labor Markets and Education." Third Way. Washington, DC. http://economics.mit.edu/files/8754.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2016a. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." NBER Working Paper 22267.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2016b. "School Quality and the Gender Gap in Educational Achievement." *American Economic Review.* 106 (5): 289–95.

Baker, Michael, and Kevin Milligan. 2013. "Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries." NBER Working Paper 18893.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116 (4): 709–45.

Becker, Gary S., William H.J. Hubbard, and Kevin M. Murphy. 2010. "Explaining the Worldwide Boom in Higher Education of Women." *Journal of Human Capital* 4 (3): 203–41.

Bertrand, Marianne, and Jessica Pan. 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics* 5 (1): 32–64.

Beyers, Jennifer M., John E. Bates, Gregory S. Pettit, and Kenneth A. Dodge. 2003. "Neighborhood Structure, Parenting Processes, and the Development of Youths' Externalizing Behaviors: A Multilevel Analysis." *American Journal of Community Psychology* 31 (1–2): 35–53.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122 (1): 409–39.

Brenøe, Anne Ardila, and Shelly Lundberg. 2018. "Gender Gaps in the Effects of Childhood Family Environment: Do They Persist into Adulthood?" *European Economic Review* 109: 42–62.

Buchmann, Claudia, and Thomas A. DiPrete. 2006. "The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement." *American Sociological Review* 71 (4): 515–41.

Cherlin, Andrew J. 2009. *The Marriage-Go-Round: The State of Marriage and the Family in America Today*. New York: Random House.

Cherlin, Andrew J. 2014. *Labor's Love Lost: The Rise and Fall of the Working-Class Family in America*. New York: Russell Sage Foundation.

Chetty, Raj, and Nathaniel Hendren. 2018a. "The Impacts of Neighborhoods on Intergenerational Mobility I: Child Exposure Effects." *Quarterly Journal of Economics* 133 (3): 1107–62.

Chetty, Raj, and Nathaniel Hendren. 2018b. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *Quarterly Journal of Economics* 133 (3): 1163–1228.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi. 2016. "Childhood Environment and Gender Gaps in Adulthood." *American Economic Review.* 106 (5): 282–88.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." NBER Working Paper 24441.

Chong, Alberto, Isabelle Cohen, Erica Field, Eduardo Nakasone, and Maximo Torero. 2016. "Iron Deficiency and Schooling Attainment in Peru." *American Economic Journal: Applied Economics* 8 (4): 222–55.

Dahl, Gordon B., and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies* 75 (4): 1085–1120.

DiPrete, Thomas A., and Claudia Buchmann. 2013. *The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools.* New York: Russell Sage Foundation.

DiPrete, Thomas A., and Jennifer L. Jennings. 2012. "Social and Behavioral Skills and the Gender Gap in Early Educational Achievement." *Social Science Research* 41 (1): 1–15.

Duckworth, Angela L., and Martin E.P. Seligman. 2005. "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents." *Psychological Science* 16 (12): 939–44.

Fan, Xiaodong, Hanming Fang, and Simen Markussen. 2015. "Mothers' Employment and Children's Educational Gender Gap." NBER Working Paper 21183.

Farrell, Albert D., Terri N. Sullivan, Layla E. Esposito, Aleta L. Meyer, and Robert F. Valois. 2005. "A Latent Growth Curve Analysis of the Structure of Aggression, Drug Use, and Delinquent Behaviors and Their Interrelations over Time in Urban and Rural Adolescents." *Journal of Research on Adolescence* 15 (2): 179–204.

Field, Erica, Omar Robles, and Maximo Torero. 2009. "Iodine Deficiency and Schooling Attainment in Tanzania." *American Economic Journal: Applied Economics* 1 (4): 140–69.

Figlio, David N., Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2013. "The Effects of Poor Neonatal Health on Children's Cognitive Development." NBER Working Paper 18846.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2014. "The Effects of Poor Neonatal Health on Children's Cognitive Development." *American Economic Review* 104 (12): 3921–55.

Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps. 2015. "Leaving Boys Behind: Gender Disparities in High Academic Achievement." *Journal of Human Resources* 50 (3): 549–79.

Gayle, George-Levi, Limor Golan, and Mehmet A. Soytas. 2015. "What Accounts for the Racial Gap in Time Allocation and Intergenerational Transmission of Human Capital?" Federal Reserve Bank of St. Louis 2015-018A.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." NBER Working Paper 12139.

Greve, Jane, Marie Louise Schultz-Nielsen, and Erdal Tekin. 2015. "Fetal Malnutrition And Academic Success: Evidence from Muslim Immigrants in Denmark." NBER Working Paper 21545.

Heckman, James J., and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics* 19 (4): 451–64

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4): 903–34.

Ingoldsby, Erin M., and Daniel S. Shaw. 2002. "Neighborhood Contextual Factors and Early-Starting Antisocial Pathways." *Clinical Child and Family Psychology Review* 5 (1): 21–55.

Karriker-Jaffe, Katherine J., Vangie A. Foshee, Susan T. Ennett, and Chirayath Suchindran. 2009. "Sex Differences in the Effects of Neighborhood Socioeconomic Disadvantage and Social Organization on Rural Adolescents' Aggression Trajectories." *American Journal of Community Psychology* 43 (3–4): 189–203.

Kearney, Melissa S., and Phillip B. Levine. 2012. "Why Is the Teen Birth Rate in the United States So High and Why Does It Matter?" *Journal of Economic Perspectives* 26 (2): 141–63.

Leventhal, Tama, and Jeanne Brooks-Gunn. 2003. "Moving to Opportunity: An Experimental Study of Neighborhood Effects on Mental Health." *American Journal of Public Health* 93 (9): 1576–82.

Lundberg, Shelly. 2017. "Father Absence and the Educational Gender Gap." IZA Discussion Paper 10814.

Lundberg, Shelly, and Robert A. Pollak. 2007. "The American Family and Family Economics." *Journal of Economic Perspectives* 21 (2): 3–26.

Lundberg, Shelly, Sabrina Wulff Pabilonia, and Jennifer Ward-Batts. 2007. "Time Allocation of Parents and Investments in Sons and Daughters." http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.619.753&rep=rep1&type=pdf.

Maccini, Sharon, and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review* 99 (3): 1006–26.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. "Internal Migration in the United States." *Journal of Economic Perspectives* 25 (3): 173–96.

Murnane, Richard J. 2013. "U.S. High School Graduation Rates: Patterns and Explanations." *Journal of Economic Literature* 51 (2): 370–422.

Norberg, K. 2004. "Partnership Status and the Human Sex Ratio at Birth." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271 (1555): 2403–10.

OECD. 2013. *Education at a Glance 2013: OECD Indicators.* OECD Publishing.

Owens, Jayanti. 2016. "Early Childhood Behavior Problems and the Gender Gap in Educational Attainment in the United States." *Sociology of Education* 89 (3): 236–58.

Prevoo, Tyas, and Bas ter Weel. 2014. "The Effect of Family Disruption on Children's Personality Development: Evidence from British Longitudinal Data." IZA Discussion Paper 8712.

Ramirez-Valles, Jesus, Marc A. Zimmerman, and Lucia Juarez. 2002. "Gender Differences of Neighborhood and Social Control Processes: A Study of the Timing of First Intercourse among Low-Achieving, Urban, African American Youth." *Youth and Society* 33 (3): 418–41.

Riphahn, Regina T., and Caroline Schwientek. 2015. "What Drives the Reversal of the Gender Education Gap? Evidence from Germany." *Applied Economics* 47 (53): 1–28.

Sampson, Robert J., Jeffrey D. Morenoff, and Stephen Raudenbush. 2005. "Social Anatomy of Racial and Ethnic Disparities in Violence." *American Journal of Public Health* 95 (2): 224–32.

Segal, Carmit. 2013. "Misbehavior, Education, and Labor Market Outcomes." *Journal of the European Economic Association* 11 (4): 743–79.

Snyder, Thomas D., and Sally A. Dillow. 2015. *Digest of Education Statistics 2013*. National Center for Education Statistics, U.S. Department of Education. Washington, DC, May.

Trivers, Robert L., and Dan E. Willard. 1973. "Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring." *Science* 179 (4068): 90–92.

Woessmann, Ludger. 2015. "Single-Parent Families and Student Achievement: An International Perspective." https://educationnext.org/files/ednext_XV_2_woessmann_unabridged.pdf.