AEA Papers and Proceedings 2021, 111: 179–183
https://doi.org/10.1257/pandp.20211124

GENDER EFFECTS IN TEACHING AND LEARNING

The Family Origin of the Math Gender Gap Is a White Affluent Phenomenon†

By Gaia Dossi, David Figlio, Paola Giuliano, and Paola Sapienza*

Differences in the gender gap in mathematics are strongly correlated with societal gender norms in the United States and across the world (Guiso et al. 2008, Nollenberger et al. 2016, Pope and Sydnor 2010). In this paper, we investigate whether the math gender gap varies by race and socioeconomic status. Using a large dataset combining information from the Florida Department of Education and birth records, we show that girls only perform worse in mathematics if they grew up in white families. In Black families, girls systematically outperform boys in mathematics, consistent with Autor et al. (2019). There are at least two nonmutually exclusive explanations for this result. On one hand, it is possible that in disadvantaged families, lower child-rearing inputs (e.g., nutrition, safety in the home, parental attention) have a disproportionate negative effect on the educational and behavioral outcomes of school-age boys relative to girls. Alternatively, biased gender norms could affect mostly girls in affluent families, because these parents have the educational and financial means to overinvest on the favored gender. This paper investigates the latter hypothesis: whether the greater gender gap in affluent white families is linked to biased educational investments in families with gender-biased norms.

Dossi et al. (2021) study the role of gender norms transmitted within the family in explaining the gender gap in mathematics. Using a variety of datasets, the authors find that maternal gender role attitudes are transmitted within the family and could help to partially explain the lower performance of girls in mathematics.

Using a similar methodology, we study whether measuring the correlation between biased family gender norms and the gender gap in mathematics could help uncover the underlying causes for the differential effect by race. Following Bharadwaj et al. (2015), Dossi et al. (2021), and Dahl and Moretti (2008), we define gender biases within the family using fertility-stopping rules and identify "boy-biased" families as those families where there are only girls except for a boy as the last born.

We find that gender role norms can explain the lower performance of girls in mathematics only in relatively affluent white families, whereas they do not apparently matter for the performance of Black girls.

Overall, our results indicate that only boy-biased white families with higher income and maternal education impact negatively girls' math achievement. This result is consistent with the hypothesis that gender-biased norms combined with greater educational and financial resources may have the perverse effect of creating a larger gap between boys and girls, whereas families with limited resources are less likely to contribute to the gender gap in mathematics, notwithstanding their gender biases. This result also helps to shed light on Fryer and Levitt's (2010) result that the biggest gender gaps are observed for white girls and for girls in the top quintile of socioeconomic status.

* Dossi: London School of Economics (email: g.g.dossi@lse.ac.uk); Figlio: School of Education and Social Policy, Northwestern University (email: figlio@ northwestern.edu); Giuliano: UCLA Anderson School of Management (email: paola.giuliano@anderson.ucla.edu); Sapienza: Kellogg School of Management, Northwestern University (email: Paola-Sapienza@kellogg.northwestern.edu). We thank seminar participants at the AEA Meetings for comments that substantially improved the paper. Opinions expressed herein do not reflect the views of the Florida Departments of Education or Health.

† Go to https://doi.org/10.1257/pandp.20211124 to visit the article page for additional materials and author disclosure statement(s).

I. Data and Variables of Interest

We make use of data from the Florida Education Data Warehouse merged with individual-level information coming from the Florida Bureau of Vital Statistics birth certificates. Our cohort is restricted to children born between 1994 and 2002, the period for which we have access to birth certificates. School records contain information on K–12 students who attended Florida public schools between the academic years 2002–2003 and 2011–2012.

Our outcome variables are standardized scores in mathematics (the Florida Comprehensive Assessment Test, or FCAT) from sixth to tenth grade.[1] Controls include children's characteristics (age in month, gender, race, and whether the child participates in a special education program) and family characteristics (whether the child is eligible for free or reduced lunch or attends a "provision 2" school), all taken from the school records. Education, marital status, and maternal age are obtained from the birth certificates together with the zip code at the time of birth.[2] Each regression also includes birth order, grade, school, and year fixed effects. More details on each variable are provided in the online Appendix.

We use fertility-stopping rules to classify families as "boy biased." We follow Bharadwaj et al. (2015), Dossi et al. (2021), and Dahl and Moretti (2008), who found evidence of parental preferences for boys over girls by showing that the number of children in the United States and Florida is significantly higher in families where the firstborn is a girl. We define "boy-biased" families with a dummy equal to one if all children are girls except for the last born and equal to zero for all the other families.[3] When comparing girls raised in gender-biased families with those who were not, we always drop the last born from our sample, since by definition the last born in a "boy-biased" family is always a boy. To rule out the possibility that the results are driven by specific dynamics related to family composition (perhaps the presence of mostly girls in the family does not allow them to learn from boys, who typically do better in mathematics), we present our regressions further restricting the sample to firstborn children. As a "placebo" group, we also look at the performance of boys in "girl-biased" families defined with a dummy equal to one if all children are boys except for the last born, who is a girl, and equal to zero otherwise.

Sample statistics for all our variables of interests are reported in Table A1 of the online Appendix.

II. Estimation Results

We start by looking at the gender gap in our sample. Girls perform worse than boys in mathematics in a regression including grade, year, and school fixed effects and a large set of individual controls (Table 1, column 1).[4] The beta coefficient is equal to -0.033. This is comparable to three-fourths of the difference between students who have a mother who is a high school dropout and those with a mother who is a high school graduate.

As a second step, we split our sample by race. Looking at the heterogeneity by race uncovers interesting differences. The gender gap is pronounced among white girls (the beta coefficient is equal to -0.048, almost double the

[1] The FCAT is the state's high-stakes criterion-referenced test. Students enrolled in public school in grades 3 through 10 are required to take the math portion every year. Students are also tested in reading, but we focus on math because of the broad-based public discussion of women and STEM. The literature has shown that the gender gap in mathematics starts appearing during junior high school (Fryer and Levitt 2010); for that reason we limit our regressions to students from sixth to tenth grade.

[2] To qualify for free or reduced lunch, the family income has to be respectively below 185 percent and 130 percent of the federal income poverty level. For details on provision 2 schools, see http://www.fns.usda.gov/school-meals/provisions-1-2-and-3. Categories for special education include mentally handicapped; orthopedically, speech, language, or visually impaired; and deaf or hard of hearing. They also include students with emotional or behavioral disabilities, autistic spectrum disorder, and other forms of serious disabilities (such as traumatic brain injuries). For maternal education, we define dummies for high school completion, some years of college, and four or more years of college. In the regressions, the excluded dummy is high school dropout mothers.

[3] For details about the construction of the sample, see the online Appendix and Dossi et al. (2021).

[4] Controls include race, age in months, birth order fixed effects, eligibility for free or reduced lunch, a dummy if the student is in a special education program, and several characteristics of the mother (age at birth, marital status, and level of education). We also control for the median income of the zip code at birth.

TABLE 1— GENDER GAP IN MATHEMATICS BY RACE

	Math score		
	All (1)	White (2)	Black (3)
Female	−0.061	−0.083	0.011
	(0.003)	(0.004)	(0.006)
Female (standardized beta)	−0.033	−0.048	0.006
Observations	703,654	489,903	154,253
R^2	0.355	0.291	0.276

Notes: This table reports OLS estimates, with robust standard errors clustered at the student and school level. The unit of observation is a student year. The data come from the Florida Department of Education (FLDOE) and the Florida Department of Health. The sample includes all students born in Florida between 1994 and 2002 enrolled in grades six to ten in a Florida public school for whom we have a mathematics test score. Sample statistics for this sample are reported in Appendix Table A1. In column 2, we restrict the sample to white students. In column 3, we restrict the sample to Black students. The dependent variable measures students' FCAT math score in a given grade (standardized with mean zero and variance one over the population for a given grade and year). In all columns, we control for median income in zip code of birth in US dollars (taken from the 1999 US census), a "free lunch" dummy variable equal to one if the student is enrolled in the free lunch program in the given academic year, a dummy for "mother married at birth" equal to one if the mother was married when the child was born, and a dummy for "special education" equal to one if the student is enrolled in the special education program in the given academic year. Column 1 includes race fixed effects. All columns include fixed effects for year, grade, and school.

coefficient of the complete sample). A different picture appears when we investigate the sample of Black girls, who instead perform better than boys (although the beta coefficient is relatively small and equal to 0.006). This evidence is consistent with results of Autor et al. (2019): boys fare comparatively worse than girls in education in disadvantaged families. The gender gap is therefore overall driven by girls growing up in white families.

As a second step, we investigate what drives the differences in mathematics performance by race. Dossi et al. (2021) demonstrate that gender biases are important in explaining the differential performance in mathematics of boys and girls. Following their approach, we classify families as "boy biased" (or "girl biased" in our placebo exercise) using fertility-stopping rules as explained above.

One interesting question is whether gender biases are relevant only in more affluent families when parents have resources or time to differentiate among their children. Fryer and Levitt (2010) show that girls fall behind boys in math relatively more in families with higher maternal education. We go further and investigate if these differences by socioeconomic status are concentrated among those families that have gender-biased beliefs about the role of women in societies.

In Table 2, we split the sample of white and Black students using two measures of socioeconomic status: eligibility for free or reduced-price lunch and maternal education. Panel A/B present the results for white and Black students, restricting the sample to firstborn girls. Column 1 reports the results for the overall sample of white students, where girls perform worse in mathematics (the beta coefficient is −0.022). When we further split the sample by eligibility for free lunch, the results appear to be driven by relatively affluent families (the coefficient is significant when we restrict the sample to families not eligible for free and reduced lunch with a beta coefficient of −0.031). The differences are similar when we split the sample by education of the mother (the beta coefficient is −0.030 for girls whose mothers have at least some college education).

These differences are not present among girls belonging to Black families, where the effect of gender bias is null and there is no significant difference across socioeconomic status.

Table 3 reports our placebo exercise where we look at the performance of boys by race and socioeconomic status. Gender biases inside the families are not correlated with the performance of boys, reducing the possibility that our measure of gender bias based on fertility-stopping rules is picking up some other omitted variables related to specific family dynamics across siblings of the same gender.

Overall, our results indicate that gender biases inside the family matter for white affluent families. In Black families, girls do perform better than boys, but the results cannot be explained by differences in gender roles transmitted by their parents. Thus, it is likely that the relative underperformance of boys in Black families is due to the fact that lower child-rearing inputs or other aspects of disadvantage have a disproportionate negative effect on boys in these families (Autor et al. 2019).

Table 2— Girls' Performance in Mathematics and Gender Biases, Heterogeneity by Race and Socioeconomic Status

	Math score				
	All families (1)	Families with FRL (2)	Families without FRL (3)	Mother attended HS (4)	Mother attended college (5)
Panel A. White, only firstborn girls					
Boy bias	−0.035	−0.013	−0.047	−0.019	−0.045
	(0.011)	(0.019)	(0.015)	(0.019)	(0.015)
Boy bias (standardized beta)	−0.022	−0.008	−0.031	−0.012	−0.030
Observations	50,402	19,223	31,179	19,062	31,340
R^2	0.297	0.310	0.232	0.313	0.229
Panel B. Black, only firstborn girls					
Boy bias	0.009	0.002	0.055	−0.031	0.015
	(0.041)	(0.047)	(0.128)	(0.070)	(0.063)
Boy bias (standardized beta)	0.005	0.001	0.036	−0.018	0.010
Observations	5,455	4,426	1,029	2,801	2,654
R^2	0.464	0.473	0.725	0.528	0.551

Notes: This table reports OLS estimates, with robust standard errors clustered at the student and school level. The unit of observation is a student year. The data come from the Florida Department of Education (FLDOE) and the Florida Department of Health. The sample includes all students born in Florida between 1994 and 2002 from a family for whom we were able to reconstruct the fertility history without any gap, and none of the children have an unknown father. From these families, we keep students enrolled in grades six to ten for whom we have a math test score. In this table, we look only at firstborn female students, excluding only children. Sample statistics for this sample are reported in online Appendix Table A2, panel A. In column 2, the sample is restricted to families with at least one child eligible for free or reduced-price lunch (FRL) in at least one year in our sample. In column 3, the sample is restricted to students from families where no child ever was eligible for FRL in any year. In column 4, the sample is restricted to students whose mother is either a high school dropout or a high school graduate (never enrolled in college). In column 5, the sample is restricted to students whose mother attended college. The dependent variable measures the student score in the FCAT in mathematics in a given grade (the score is standardized with mean zero and variance one over the population for a given grade and year). "Boy bias" is a dummy equal to one if the last-born child in the family is a boy and all the older children are girls, equal to zero otherwise. The set of controls is identical to the one described in Table 1. All columns include fixed effects for year, grade, and school.

III. Conclusions

Previous research has shown that norms around the role of women in society could help explain the gender gap in mathematics and that these norms could be transmitted within the family.

Using data from the Florida Department of Education combined with birth certificates, we uncover important heterogeneity in the transmission of gender biases within the family. Gender-biased attitudes (proxied in this paper by fertility-stopping rules) are correlated with lower performance in mathematics only in white affluent families. Girls do better in Black families, where the results do not appear to be driven by a different transmission of gender norms.

REFERENCES

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." *American Economic Journal: Applied Economics* 11 (3): 338–81.

Bharadwaj, Prashant, Gordon B. Dahl, and Ketki Sheth. 2015. "Gender Discrimination in the Family." In *The Economics of the Family*, Vol. 1, edited by Esther Redmount, 237–66. Santa Barbara, CA: ABC-CLIO.

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

Dossi, Gaia, David N. Figlio, Paola Giuliano and Paola Sapienza. 2021. "Born in the Family:

Table 3—Boys' Performance in Mathematics and Gender Biases, Heterogeneity by Race and Socioeconomic Status

	Math score				
	All families (1)	Families with FRL (2)	Families without FRL (3)	Mother attended HS (4)	Mother attended college (5)
Panel A. White, only firstborn boys					
Girl bias	−0.004	0.003	−0.012	−0.025	0.008
	(0.011)	(0.019)	(0.014)	(0.019)	(0.014)
Girl bias (standardized beta)	−0.002	0.002	−0.007	−0.015	0.005
Observations	52,737	19,795	32,942	19,815	32,922
R^2	0.329	0.357	0.260	0.340	0.269
Panel B. Black, only firstborn boys					
Girl bias	0.014	0.020	−0.079	0.114	−0.027
	(0.049)	(0.058)	(0.118)	(0.076)	(0.075)
Girl bias (standardized beta)	0.008	0.010	−0.046	0.059	−0.015
Observations	5,002	3,995	1,007	2,506	2,496
R^2	0.478	0.488	0.710	0.572	0.545

Notes: This table reports OLS estimates, with robust standard errors clustered at the student and school level. The unit of observation is a student year. The data come from the Florida Department of Education (FLDOE) and the Florida Department of Health. The sample includes all students born in Florida between 1994 and 2002 from a family for whom we were able to reconstruct the fertility history without any gap, and none of the children have an unknown father. From these families, we keep students enrolled in grades six to ten for whom we have a mathematics test score. In this table, we look only at firstborn male students, excluding only children. Sample statistics for this sample are reported in online Appendix Table A2, Panel B. Columns 1 to 5 are defined in the same way as in Table 2. "Girl bias" is a dummy equal to one if the last-born child in the family is a girl and all the older children are boys, equal to 0 otherwise. The dependent variable and the controls are identical to the ones described in Table 2. All columns include fixed effects for year, grade, and school.

Preferences for Boys and the Gender Gap in Math." *Journal of Economic Behavior & Organization* 183: 175–88.

Fryer, Roland G., Jr., and Steven D. Levitt. 2010. "An Empirical Analysis of the Gender Gap in Mathematics." *American Economic Journal: Applied Economics* 2 (2): 210–40.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. 2008. "Culture, Gender, and Math." *Science* 320 (5880): 1164–65.

Nollenberger, Natalia, Núria Rodriguez-Planas and Almudena Sevilla. 2016. "The Math Gender Gap: The Role of Culture." *American Economic Review* 106 (5): 257–61.

Pope, Devin G., and Justin R. Sydnor. 2010. "Geographic Variation in the Gender Differences in Test Scores." *Journal of Economic Perspectives* 24 (2): 95–108.