ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 23-707

When Girls Outperform Boys: The Gender Gap in High School Math Grades

Lauren Sartain
The University of North Carolina at Chapel Hill

Silvana Freire
New York University

John Q. Easton
University of Chicago

Briana Diaz
University of Chicago

Across an array of educational outcomes, evidence suggests that girls outperform boys on average. For example, in Chicago, ninth-grade girls earn math GPAs that are 0.29 points higher than boys on average. This paper examines explanations for this gap, such as girl-boy differences in academic preparation, behaviors and habits, and experiences in math classes. After accounting for these factors, the gender gap in math grades persists. We, then, examine the classroom-level conditions that reduce the gender gap in grades. The gap is smaller in more advanced courses like honors classes and geometry. Further, boys perform more similarly to girls in classes with male teachers. These findings highlight classroom conditions that are more conducive to the academic success of boys.

VERSION: January 2023

Suggested citation: Sartain, Lauren, Silvana Freire, John Q. Easton, and Briana Diaz. (2023). When Girls Outperform Boys: The Gender Gap in High School Math Grades. (EdWorkingPaper: 23-707). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/b0py-tz14

When Girls Outperform Boys: The Gender Gap in High School Math Grades

By Lauren Sartain, Silvana Freire, John Q. Easton, and Briana Diaz

June 2021

Abstract

Across an array of educational outcomes, evidence suggests that girls outperform boys on average. For example, in Chicago, ninth-grade girls earn math GPAs that are 0.29 points higher than boys on average. This paper examines explanations for this gap, such as girl-boy differences in academic preparation, behaviors and habits, and experiences in math classes. After accounting for these factors, the gender gap in math grades persists. We, then, examine the classroom-level conditions that reduce the gender gap in grades. The gap is smaller in more advanced courses like honors classes and geometry. Further, boys perform more similarly to girls in classes with male teachers. These findings highlight classroom conditions that are more conducive to the academic success of boys.

Keywords: gender, math, grades, high school.

Acknowledgments: The authors gratefully acknowledge the *William & Flora Hewlett Foundation* and the UChicago Consortium Investor Council for funding this research, and partners from Chicago Public Schools for providing access to the data and helping us better understand it. This paper benefited from feedback from discussion with UChicago Consortium researchers and presentation at the American Educational Research Association and the Association for Education Finance and Policy annual conferences. The authors thank Naureen Kheraj for her contributions to the analysis and the initial version of this paper.

I. Introduction

Female students outperform male students in different measures of academic performance throughout middle school and high school (Gibb, Fergusson & Horwood, 2008; Voyer & Voyer, 2014). Although for a long time, studies reported a male advantage on standardized tests in the areas of math and science, research has shown that the male-female gap in mathematics has narrowed, especially for older students (Fahle & Reardon, 2018; Hyde, Lindberg, Linn, Ellis & Williams, 2008; Lindberg, Hyde, Petersen & Linn, 2010; Reardon, Fahle, Kalogrides, Podolsky & Zarate, 2018). In addition, female students are also more likely to graduate from high school (Allensworth, Healey, Gwynne, & Crespin, 2016; Green & Winters, 2006; Stetser & Stillwell, 2014), as well as to attend college and ultimately attain a degree than their male peers (DiPrete & Buchmann, 2013; Goldin, Katz & Kuziemko, 2006; Jacob, 2002).

When it comes to course grades, on average, females also earn higher high school GPAs than males (Authors, 2017; DiPrete &Buchmann, 2013; Perkins, Kleiner, Roey, & Brown, 2004; Voyer & Voyer, 2014). Although the female-male differences in grades are largest in language courses, girls perform better than boys, on average, across different school subjects, including STEM-related subjects such as math, science, and computer science (Authors, 2020; Duckworth & Seligman, 2006; O'Dea, Lagisz, Jennions & Nakagawa, 2018; Voyer & Voyer, 2014). Even among students who take more advanced coursework, girls still show a GPA advantage compared to boys of 0.20 grade-points (DiPrete & Buchanann, 2013).

The fact that boys perform relatively worse than girls in their high school coursework is troubling, as a growing body of research points to the importance of grades as a reliable indicator that strongly predicts future educational outcomes. Grades are not only a strong predictor of educational attainment, they also predict post-secondary outcomes like college enrollment, persistence, and graduation better than test scores do (Authors, 2017; Allensworth & Clark, 2020; Bowen, Chingos & McPherson, 2009; Jacob, 2002). Ultimately, the observed gender difference in grades suggests that girls are generally better positioned to graduate from high school and to enroll and succeed in college. Yet, there are unanswered questions about the reasons for these gender differences in grades, leaving policy and practice without evidence-based guidance on how to create more supportive learning environments for all students.

In this paper, we examine the gender differences in teacher-assigned grades in ninth-grade mathematics classes in Chicago Public Schools (CPS). CPS is the third-largest school district in the country, serving students with varying academic performance levels in a wide range of school settings. This study focuses on math because it is a critical course for ninth graders and can serve as a gatekeeper for students, as failure rates in algebra courses are usually high and lead to students being off track and less likely to graduate from high school (Heppen et al., 2016). Using a unique set of administrative and survey data sources for nearly 10,000 ninth graders in CPS from the 2016-17 and 2017-18 school years, we examine three potential student-level explanatory factors for the gender gap in math grades: academic preparation (i.e., 8th-grade test scores), behaviors and habits (i.e., attendance, suspensions, academic effort, and persistence), and self-reports of math classroom experiences. Being able to account for these factors is important since prior research has shown that teacher-assigned grades are not only a reflection of cognitive skills but also capture different non-cognitive dimensions of students' learning (see Bowers, 2011 and Brookhart et al., 2016). In that sense, gender differences in key academic and behavioral measures may help explain the observed gender gap in math GPA. At the classroom level, we consider additional relevant factors such as the type of math course (e.g., algebra, geometry) and whether it was an honors-level class, as well as the classroom composition and teacher characteristics. Specifically, we answer the following research questions:

1. What is the extent of the gender gap in students' ninth-grade math grades? How does the grade gap compare to gender differences in other student-level factors like academic preparation, behaviors and habits, and reported math classroom experiences?
2. How do key academic and behavioral measures and student reports of their experiences in math class contribute to the gender difference in ninth-grade math GPA?
3. Are there classroom settings where the gender gap in math grades is diminished, or even reversed?

We find that girls received ninth-grade math grades that are, on average, about 0.29 GPA points higher than boys. We also see that girls outperform boys in many other dimensions: their average incoming reading test scores are higher, and they typically have lower rates of disciplinary infractions in ninth grade and have higher reports of academic effort. Girls are also, on average, more likely to report that their math teachers challenge them academically while also providing clear expectations and support for learning. Because these other factors are also strong predictors of grades, we hypothesized that accounting for these gender differences would narrow or eliminate the gender gap in math grades. However, after accounting for all of these differences in academic preparation, behaviors and habits, and classroom experiences, the estimated gender gap in grades remained at about 0.25 GPA points, large in magnitude and little changed from the unadjusted gap. We, then, consider if there are classroom-level factors that might lessen the gender gap in math grades and point to examples of classroom environments where boys are better set up to succeed academically. Our classroom-level analysis points to some contexts where the gap is smaller: in advanced math courses like honors classes and geometry and in classes with male teachers. These findings highlight classroom conditions that may be more conducive to the academic performance of boys and should be explored further to provide more insights to practitioners.

II. How Academic and Non-Academic Factors Relate to Teacher-Assigned Grades

Recent research recognizes teacher-assigned grades as a valid and comprehensive measure of multiple dimensions of student learning (see Bowers, 2011 and Brookhart et al., 2016). Grades capture students' content knowledge and also reflect day-to-day skills and behaviors that teachers find important for academic success (Authors, 2017; Farrington, Roderick, Allensworth, Nagaoka, Keyes, Johnson & Beechum, 2012). In a rapidly changing world, students not only need to demonstrate a clear grasp of core academic content but must apply that understanding to the outside world while simultaneously developing a range of soft skills and competencies. These skills and knowledge known as "deeper learning" will enable them to successfully participate in the 21st-century workplace and civic life (William and Flora Hewlett Foundation, 2013). Since grades reflect, at least in part, these important skills, gender differences in terms of students' academic preparation, behaviors, and experiences in their classrooms may explain why boys tend to get lower grades. In this section, we examine some of these key academic and behavioral measures, as well as features of classroom environments, that might be contributing to the gender gap researchers and practitioners observe in students' final grades.

A. Gender Differences in Academic Preparation

As is widely known, grades are correlated with test scores, and therefore they in part reflect mastery of academic content knowledge like what is assessed by standardized tests. A comprehensive review of research on teacher-assigned grades reveals that grades typically correlate about 0.5 with standardized measures of achievement (Brookhart et al., 2016). In addition, ninth-grade GPA has been found to be significantly predictive of students' test score gains in later grades (Authors, 2017). In this sense, any gender differences in students' test achievement and academic preparation may explain part of the differences in grades. However, recent research shows that the math test score gap is small or non-existent (Fahle & Reardon, 2018; Hyde et al., 2008; Lindberg et al., 2010). Because girls and boys perform similarly on standardized math tests, the observed gender differences in teacher-assigned grades is likely associated with differences in other student behaviors such as attitude towards study, homework completion, and classroom conduct among other measures which are not fully captured by performance on standardized tests. Additionally, any gender differences in the ways that students experience their classroom environments, including perceptions of teacher support and expectations, and academic challenge, may also translate to differences in grades.

B. Gender Differences in Behaviors and Habits

Many researchers have turned to analyzing social and behavioral skills such as student effort, engagement, motivation, persistence, and self-regulatory behavior, in order to understand why girls are generally outperforming boys in their classes (DiPrete & Jennings, 2012; Downey & Vogt Yuan, 2005; Duckworth & Seligman, 2006; Lam et al., 2012). In particular, school engagement has received much attention as a malleable dimension of student learning that can directly impact academic achievement. Although it is often regarded as a multidimensional construct, the behavioral aspect of it relates to students' participation in learning and academic tasks (i.e. effort, persistence, attention, contributing to class discussion) and the display of positive conduct such as following rules and completing homework. Similarly, the cognitive and motivational component of engagement stresses students' self-regulatory skills, effort and investment toward explicit learning goals (Fredricks, Blumenfeld & Paris, 2014). These behaviors are considered crucial for achieving positive academic outcomes.

Several studies have found important female-male differences when it comes to these types of behaviors and academic habits. For instance, girls may appear as more motivated and engaged with school and schoolwork than boys. A study on academic engagement and performance across 12 countries, including the U.S., found that girls reported higher levels of school engagement than boys; which partially explained higher academic performance ratings by their teachers (Lam et al., 2012). Similarly, a study in Canada showed that boys in higher grades were more likely to be academically disengaged and less likely to perceive social support from their friends and peers (Wilcox, McQuay, Blackstaffe, Perry & Hawe, 2018).

Boys' school disengagement may also be evidenced by their higher suspension rates compared to girls. On average, male students have higher rates of school disciplinary problems, such as skipping class or getting into fights, get suspended or expelled at higher rates than girls do (DiPrete & Buchmann, 2013; Finn & Servoss, 2014; Skiba, Michael, Nardo & Peterson, 2002), and are more likely to be retained in the ninth grade (Jacob, 2002; Warren, Hoffman & Andrew, 2014). Teachers also seem to perceive boys as being more disruptive in the classroom. For instance, Gibb, Fergusson and Horwood (2008) found that teachers rated boys as more prone to inattentive, restless and oppositional classroom behaviors, which seemed to explain a substantial portion of the gender differences in educational achievement. These differences could then translate into differences in grades between female and male students to the extent that grades reflect these types of behaviors.

Other research provides further evidence on the role that students' behaviors and social skills play in their academic outcomes throughout their school careers. Using data from the Early Child Longitudinal Study–Kindergarten survey, DiPrete and Jennings (2012) found that in fact gender differences in the acquisition of important social and behavioral skills explained a substantial portion of the gender gaps on academic outcomes in elementary school. Likewise, Downey and Vogt Yuan (2005) study on middle grade and high school students concluded that a major reason for the observed gender differences in grades was boys' poorer classroom behavior in terms of effort and disruptiveness.

At the same time, the deeper learning literature maintains that academic engagement is associated with students' academic mindsets. These motivational components influence how much effort and persistence students' put into their work, to problem-solve and overcome difficulties (William and Flora Hewlett Foundation, 2013). Research shows that the level of effort that middle school and high school male and female students put into schoolwork differs, and these differences are related to the gender gap in academic performance. Studies based on students' time-logs and survey reports found that girls complete more out-of-school homework and spend more time studying than their male peers (DiPrete & Buchman, 2013; Gershenson & Holt, 2015). According to the National Household Education Surveys Program (NHES) 2016 data, high school girls are 6 percentage points more likely to do homework outside of school and report spending on average 2.2 additional hours per week doing homework compared to boys (U.S. Department of Education, National Center for Education Statistics, 2017).

Although most of this evidence suggests that girls are often more academically motivated than their male peers, a recent study revealed that self-control appears as a more important factor than motivation in explaining female's higher grades. Teachers', parents' and students' self-reports indicated that girls displayed more self-control and motivation than boys. However, further statistical analyses found that only self-control mediated the relationship between gender and academic performance, measured as improvements and overall final GPA (Duckworth, Shulman, Mastronarde, Patrick, Zhang, & Druckman, 2015). Findings from this and prior studies on self-regulatory behaviors and academic performance (see Duckworth & Seligman, 2006; Gil Gomes Carvalho, 2016) show that girls not only exhibit a greater ability to control impulses and persist in the service of a higher goal, but that this trait explains part of female's higher grades across subjects.

C. Gender Differences in Classroom Experiences

Additionally, students' classroom experiences and the environments in which they are immersed are likely to contribute to their academic performance. However, males and females may perceive and experience their learning environments differently even within the same classroom. In this sense, certain classroom conditions may be more conducive to the academic success of girls than boys. For instance, literature shows that the degree to which teachers provide clear expectations and support students into meeting the learning goals influences their academic engagement and performance (Klem & Connell, 2004; Roorda, Koomen, Spilt & Oort, 2011). Some studies show that boys are more likely to report lower levels of teacher involvement and support than their female peers (Gherasim, Butnaru & Mairean, 2012; Lietaert, Roorda, Laevers, Verschueren & De Fraine, 2015), while other studies reported no gender differences or female-favorable gaps regarding students' perceptions of classroom support (Tennant, Demaray, Malecki, Terry, Clary & Elzinga, 2014; Samuelsson & Samuelsson, 2016). From an academic risk perspective, a supportive learning environment might be of greater importance for boys given their higher risk of school failure compared to girls (Hamre & Pianta, 2001). A meta-analysis based on 99 studies found that affective teacher–student relationships were more important for boys' than girls' school engagement (Roorda et al., 2011). To the extent that supportive classroom environments promote student engagement and ultimately influence academic performance, gender differences in the perceived levels of teacher support and the clarity of teacher expectations might contribute to the female-male differences observed in grades. If girls are in fact feeling more supported than boys, we could expect these different experiences to contribute to the GPA gap.

Similarly, challenging instructional classroom environments where teachers provide activities of higher cognitive demand and promote critical thinking will support and enable deeper learning among students (William and Flora Hewlett Foundation, 2013). Yet again, research shows that boys and girls perceive academic demands and approach their learning objectives differently. For example, a study on math classroom environments and students' achievement goals showed that girls had goals that were more mastery-oriented (focus on the developing their own competence) and less performance-avoidance oriented (focus on avoiding appearing incompetent) than boys (Gherasim et al., 2012). Students' achievement goals can be influenced by their perceptions of teacher's expectations and academic demands. A study on students' perceived academic press and achievement goals in math found that while perceptions about academic demands and challenge were similar across demographic groups, at higher levels of perceived press girls were less reluctant than boys to seek help when experiencing greater demands (Middleton & Midgley, 2002). Based on these findings we would expect girls to interpret teacher's demands in their classes as a driver to put more effort toward mastering the academic content and achieving the learning goals, and to be willing to resort to different strategies in order to meet those academic demands.

Finally, teacher expectations and perceptions of students' abilities may also play an important role in students' engagement with schoolwork and academic performance. Riegle-Crumb and Humphries (2012) examined teachers' conditional biases when perceiving students' ability by gender and race/ethnicity. Their findings suggest teachers were more likely to believe math was easier for white boys than it is for white girls. While this belief held true across different course levels examined, in average classroom contexts white females were more likely to be judged as being less prepared for the class difficulty level than boys. This pattern was not found in the advanced classes, possibly suggesting that teacher stereotypes are more salient when there's more ambiguity regarding students' skills. It may be the case that in higher-level courses, teachers may perceive that students have already completed the demanding requisites to participate in that class and are equipped with the academic skills expected at that level.

But teachers also hold beliefs about students' expected classroom behaviors and academic habits, which are also likely to differ for boys and girls. For example, teachers generally perceive boys' class behaviors more negatively and less task-oriented than girls. However, Mullola et al. (2012) found that those perceptions differed between female and male teachers. Male teachers were more likely to perceive boys more positively and more capable in educational competence and teachability than female teachers do. This goes in line with Dee's (2007) findings, where same-gender teacher-student match was related to higher achievement, more positive teacher perceptions of student performance and student engagement with the teacher's subject for both girls and boys. For example, when girls were taught by a female teacher, they were less likely to report being disengaged in class (i.e. reporting they did not look forward to a subject or that they were afraid to ask questions). These findings provide evidence of the influence that teacher characteristics such as gender have on the student expectations but also on the students' own perceptions regarding their abilities and academic engagement.

D. Contribution

As previously noted, many researchers have not only established the importance of grades as predictors of academic and lifelong success but also either suggested or demonstrated that this measure is influenced by myriad factors in addition to cognitive learning. With the rich longitudinal data archive from Chicago Public Schools containing behavioral indicators, as well as students' reports on non-cognitive constructs like the ones previewed, we intend to determine the extent of these relationships and whether the gender differences on these factors are related to the higher grades that girls earn in their math courses. This paper adds to the current literature by examining multiple factors. Past research typically focused on a single explanatory factor, such as the relationship between academic skills and grades or self-reports of study habits and grades. We examined multiple factors in order to understand better how differences between students' incoming academic preparation and their behaviors and academic habits— as well as students' reports of their experiences in their classroom contribute to the gender gap in math grades. We can account for students' perceptions about the level of support from their teachers, the depth of math instruction, and the amount of perceived academic challenge. In addition to these student-level factors, we can investigate classroom-level settings where the female-male grade gap might be diminished.

III. Data and Research Methodology

In this section we describe the analytic samples we use to answer the research questions, as well as the rich administrative and survey data from CPS. We, then, discuss the models we use to understand the relationship between key academic and behavioral measures and the gender gap in math GPA.

A. Analytic Sample

Our paper consists of two sets of analysis differentiated by their unit of analysis: 1) at the student level, and 2) at the classroom level. First, for the student-level analysis, our sample includes students who 1) enrolled in 9th grade[1] for the first time in the fall of the 2016-17 and 2017-18 school years in a district-run high school (that is, non-charter); 2) had math grades during that school year; and 3) responded to survey items about their math class (N = 9,585 students). Table 1 displays descriptive statistics for all first-time ninth graders with math grades and for students in the analytic sample. Overall, the analytic sample is demographically similar to all first-time CPS ninth graders. The sample is nearly equally split between girls (51 percent of students) and boys (49 percent of students).[2] The analytic sample is nearly one-third Black students and one-half Latinx students. Students in the analytic sample are more likely to live in Census blocks with higher levels of employment and household income (high-SES), relative to other students in the district. Nine percent of students in the sample classified for special education services and eight percent were English-Language Learners. Finally, students in the sample are relatively higher performing than the 9th-grade student population in CPS with slightly higher incoming test scores and higher math GPAs in 9th grade, on average.

Then, for the classroom-level analyses, we look at explanatory factors of the gender gap in final semester grades in 1,281 math classes (see Table 2). We excluded classes where fewer than three ninth-grade girls and three ninth-grade boys enrolled (the average class size in our analytic sample was 22 students). Most classes were algebra classes (81 percent), which is the typical class that CPS ninth graders enroll in, followed by geometry (14 percent) for students who completed algebra prior to ninth grade, and fewer advanced (4 percent) and elective (1 percent) math classes. Forty percent of math classes in the sample were taught at the honors level. We also consider the teacher's gender: 44 percent of the classes were taught by male teachers compared to 54 percent taught by a female teacher (in 2 percent of the classes teacher demographics were missing). Most classes were taught by white teachers (53 percent), followed by 18 percent of classes where teachers identified as Latinx, and 14 percent where teachers identified as Black. The typical teacher also had about 11 years of teaching experience in CPS.

B. Data and Variables

This study uses administrative data from CPS, including information on students' demographic characteristics, their academic performance, disciplinary infractions, attendance, and high school transcripts. We also use personnel records, which include teachers' demographic information. Additionally, our analyses included students' self-reports on their own academic behaviors, as well as their reports of the instructional environment in their 9th-grade math classes. Students' reports were collected through the 5Essentials survey administered by CPS to all students in the spring of 2016-17 and 2017-18 school years.

Transcript data. Our outcome of interest is students' math GPA in ninth grade. CPS provided high school transcript data, which shows each class a student takes (i.e., level, type), their final grade for each class, and the teacher of record. We calculate students' math GPA as the average of final grades earned in their ninth-grade math classes.[3] We do not use special weighting for advanced or honors classes, though we note that weighted and unweighted GPAs are highly correlated with each other. In our regression models, we included an indicator to identify students who took math honors classes and those who took at least one geometry class.

The transcript data also allows us to place students into classrooms and to link students to their math teacher.

Academic preparation. In eighth grade, all CPS students take the Northwest Evaluation Association's (NWEA) MAP exam, which we use as a measure of students' incoming academic achievement. We averaged students' eighth-grade NWEA math and reading test scores [4] and standardized within a school year to have a mean of 0 and a standard deviation of 1. We imputed eight grade test scores for the nine percent of the analytic sample with missing values.[5] Most of these students were new to CPS in ninth grade.

Student behaviors. Student behaviors measured in the administrative data include daily school attendance and disciplinary infractions. We calculated the attendance rate as the proportion of the days enrolled that a student was present in school. Disciplinary infraction data includes the number of suspensions a student had during the school year. We use a binary indicator for whether the student had one or more out-of-school suspensions.

Self-reports of behaviors and habits. Every spring CPS in partnership with the University of Chicago administers a survey to all students in grades 6-12 called the 5Essentials survey.[6] The goal of the survey is to provide school leaders and the general public with data on school culture and climate. The survey also includes items about students' own academic behaviors and habits. Guided by the 5Essentials conceptual framework, researchers at the UChicago Consortium on School Research created measures (scales) of important constructs. They used Rasch analysis to combine student responses to specific items on the survey into measures. These have been carefully examined for reliability and internal validity. Appendix Table 1 includes a complete list of survey items included in the measures we use in this paper.

We use two measures of students' academic behaviors from the survey data:

- *Academic Effort* measures students' attitudes around studying and schoolwork and their ability to manage their time and prioritize their schoolwork. This measure includes items such as "I set aside time to do my homework and study" or "I try to do well on my schoolwork even when it isn't interesting to me".
- *Persistence* measures whether students sustain interest and effort toward long-term goals. This measure is a modified version of the perseverance items from the shortened Grit scale (see Duckworth & Quinn, 2009; Duckworth, Peterson, Matthews & Kelly, 2007). It includes items such as "I continue steadily toward my goals" or "I finish whatever I begin".

Students' reports on their experiences in math classes. We also use the 5Essentials survey data to characterize students' experiences in their math classes. About one-third of students who took the survey were randomly assigned to complete questions about experiences in their math class. Our analyses included measures on student self-reports about their math instructional experience and their math class environment:[7]

- *Higher-order Math Instruction* measures whether students performed activities that required higher-order thinking skills such as critical thinking and problem-solving in their math classes. Students answered how often they did activities such as "Apply math to situations in life outside of school", "Explain how you solved a problem to the class" or "Solve a problem with multiple steps that takes more than 20 minutes".
- *Math Class Environment.* Students answered items about their math class instruction and teacher. Specifically, the survey captures student perceptions about the level of academic support, concern, and clarity from their math teachers,

as well as whether students perceived their math class as challenging with teachers pushing all students toward high levels of academic achievement (see Appendix Table 1 for a list of items included in these scales). Based on the interconnectedness of these various constructs, we used Principal Components Analysis (PCA) to group them into one single index of Class Environment (α = 0.86).[8] Because the PCA suggested applying similar weights to each of the constructs, the index is the average of the corresponding survey measures standardized within year to have a mean of 0 and a standard deviation of 1.

Student background characteristics. Data on student demographic information (gender, race/ethnicity, special education status, ELL designation, and temporary living situation–*homelessness*–status) and school enrollment came from district administrative data sets. Student socioeconomic status (SES) was measured using a poverty concentration index comprised of American Community Survey (ACS) data on the percentage of adult males employed and the percentage of families with incomes above the poverty line at the Census block group where students lived.[9] We use the ACS data to describe SES because most students in CPS qualify for free/reduced-price lunch, so the ACS data provide more nuanced detail about students' neighborhood contexts.

Math Class Characteristics. For the classroom-level analysis, we aggregate student information to the classroom-level to characterize the class composition (i.e. proportion of male students, average 8th-grade test scores). We also include indicators that identify the class level (honors or regular) and type (algebra, geometry, advanced math, or electives). The transcript data allows us to identify the teacher of record. From personnel records, we match teacher characteristics such as their gender, race/ethnicity, and years of experience (based on their hired date in CPS).

Analysis

Student-Level Factors and the Gender Gap in Math Grades. The first research question explores how gender differences in academic and behavioral factors, as well as classroom experiences, may influence the gender gap in math grades. We start by documenting the extent of the gender gap in grades and in other relevant measures of student performance, like incoming test scores, academic habits, and the ways in which students perceive teacher support and rigor in their math classes. Then, to understand more about how differences in these various factors explain the gender gap in math GPA, we estimate a series of OLS regression models, accounting for students' background characteristics and incoming achievement, student academic behaviors, and their perceptions of their math class environment.

Specifically, we estimate variations of the following fully specified regression model, accounting for school-by-year fixed effects:

$$MathGPA_{its} = \alpha + \beta Male_i + \gamma Background_i + \delta AcademicPrep_{its} + \varphi Behaviors_{its} + \tau ClassExperience_{its} + \theta_{ts} + \varepsilon_{its} \quad (1)$$

where the outcome is math GPA (in grade-point units, from 0.0 to 4.0) for student *i* from year *t* who attends school *s*. We start by estimating the simplest model to predict math GPA with a single variable: student gender. Then we build the model by adding, first, a vector of students' background characteristics, followed by students' incoming achievement, academic behaviors, and students' reports about their math class experience (each of these vectors are described above in the data description). To further approximate students' experience, the *Math Class* vector also includes an indicator variable that equals 1 if the student took at least one geometry class and an indicator variable that equals 1 if the student took at least one honors-level math course. All of the model specifications include a school-by-year fixed effect (θ_{cs}) to account for patterns of increasing grades over time and for idiosyncratic school-specific grading policies and other institutional characteristics that do not vary inside a school within a specific year. Thus, by holding time-invariant characteristics of the school-by-year constant, we compare students' math grades in the same school and cohort to each other. Standard errors are clustered at the school level to account for any correlation in unobservable factors across students who attend the same school. Ultimately, we are interested in how the estimate of $\beta_1$changes with the addition of various student characteristics and experiences.

We investigate the possibility that the relationship between gender and math GPA varies based on students' position in the GPA distribution. For example, the gender gap may be smaller or non-existent for students with high math grades. To examine these patterns, using the full model specification shown in equation (1), we estimate a quantile regression model that predicts students' math GPA at the 10th, 25th, 50th, 75th, and 90th percentiles. This approach allows us to examine the magnitude of the gender gap at different points in the performance distribution.

Classroom-Level Factors and the Gender Gap in Math Grades. With the second research question, we dive deeper into classroom conditions that may widen or lessen the gender gap in math GPA. We focus on examining the gender difference in math final grades at the classroom level to better characterize the classrooms where boys and girls are performed similarly. This analysis includes 1,281 math classes where students in the analytic sample enrolled during the spring of their ninth-grade year. For this analysis, we calculated the gender difference in grades as the difference between girls' average final grade and boys' average final grade in the classroom (see equation 2):

$$GenderDiffinGPA_{cts} = \overline{MathGPA}_{cts}^{male=0} - \overline{MathGPA}_{cts}^{male=1} \quad (2)$$

Using OLS regression, we estimated the following model:

$$GenderDiffinGPA_{cts} = \alpha + \beta ClassComposition_{cts} + \gamma CourseType\&Level_{cts} + \delta TeacherCharacteristics_{cts} + \theta_{ts} + \varepsilon_{cts} \quad (3)$$

where the outcome is the girl-boy difference in the average final grade (in grade-point units) for class *c* in school *s* in year *t*. As predictor variables, we consider the classroom composition based on the proportion of male students and the average classroom 8th grade test scores, the class type and level, and teacher's characteristics (gender, race/ethnicity, and years of experience). We estimate each one of these vectors independently and, finally, a fully specified version that includes all these measures simultaneously. All models include school fixed effects to account for institutional characteristics and policies that did not vary over time.

IV. Results

In this section, we first describe the gender gap in ninth-grade math grades and compare it to gender differences in other cognitive and behavioral student outcomes. Then, we analyze how much of the gender difference in math grades at the student level can be explained by students' background characteristics, behaviors, and their reported experiences in math classes. Finally, we explore the characteristics of classrooms that are associated with reductions in the gender gap in math GPA.

A. Gender differences in students' ninth-grade math grades and other measures

To answer our first research question, we start by comparing girl-boy differences in math GPA, as well as other key student-level factors like academic preparation, behaviors and habits, and reported math classroom experiences. Across both cohorts, girls earned higher math grades than boys with an average math GPA of 2.78 relative to 2.49. This represents a statistically significant difference of 0.26 standard deviation units (see Figure 1). Figure 1 also shows standardized differences in other key academic and behavioral measures between girls and boys (Appendix Table 2 shows the mean and standard deviation for these variables by gender, and Appendix Table 3 includes the correlations of the variables.) Notably, there are not significant differences in incoming math test scores by gender.[10] However, consistent with the literature, we see that, on average, girls outperformed boys in reading with a standardized difference of 0.20.

Turning to behaviors and habits, we do not find significant differences between girls' and boys' average attendance rates in the sample (93.6 average percent attendance rate versus 93.3 percent—a difference of about one half of a school day). However, we do see larger differences in terms of disciplinary infractions. On average, girls in the sample received fewer out-of-school suspensions than boys during ninth grade (3.9 percent of girls versus 6.1 percent of boys). This represents a difference of 0.10 standard deviation units. In addition, girls generally report much higher levels of academic effort relative to boys (a standardized difference of 0.17), though self-reports of persistence did not differ significantly by gender.

To examine differences between girls' and boys' perceptions about their math classroom experiences, we rely on students' survey reports about their math class. The higher-order math instruction measure includes items asking students about the kinds of assignments and activities teachers ask them to complete. Perhaps not surprisingly, we do not see a statistically significant difference in reports of higher-order math instruction between boys and girls. However, considering the classroom environment index that includes items related to teacher expectations, instructional clarity, and academic challenge, on average, girls report higher levels than boys do (a standardized difference of 0.05).

B. Explaining gender differences in ninth-grade math grades with student-level factors

Our second research question inquires to what extent do these key academic and behavioral measures and student reports of their class experience contribute to the gender difference in ninth-grade math GPA. While the difference in math grades favoring girls is the largest we see in all of these factors by gender, we do find that girls are suspended less often than boys, report higher levels of academic effort, and perceive that their math teachers challenge them academically while also providing clear expectations and support for learning than do boys. Based on these differences between girls and boys, and based on the literature, we now turn to exploring how much of the difference in math grades are explained by differences in these other factors. Using a series of regression analyses, we estimate how academic performance, behaviors and habits, and classroom experiences moderate the relationship between ninth-grade GPA and gender. Table 3 presents the coefficients estimated in each model specification.

In Model 1, we predict ninth-grade math grades for boys unconditional of other variables, except school-by-year fixed effects. The male coefficient is -0.25 grade points, which is very close to the observed difference without the school-by-year adjustment. Subsequently, in Model 2, we also consider prior academic performance (as measured by 8th-grade math test scores). We also include a vector of student demographics, which are in all subsequent specifications. There is a strong and statistically significant relationship between prior achievement and math GPA. A standard deviation increase in incoming math test scores corresponds to a 0.53-grade-point increase in math GPA in ninth grade. However, the male coefficient (-0.28 grade points) is similar to that in Model 1, suggesting that differences in incoming test scores does not explain the gender difference in GPA.

In Model 3, we account for student behaviors and habits, specifically ninth-grade attendance, out-of-school suspensions, and their self-reports of academic effort and persistence.

We find that all of these variables are statistically significant predictors of math GPA in the expected direction. For example, students who received at least one out-of-school suspension have GPAs that are 0.34 grade points lower, on average, than students who were not suspended. Similarly, students with higher reports of academic effort and persistence also receive higher math grades. However, the estimated coefficient for male students is almost identical to that in Model 1 (-0.24 grade points). As such, the observed gender differences in academic effort and persistence do not explain the gender gap in GPA.

Model 4 adds the students' reports of math classroom experiences (higher-order math instruction and classroom environment), as well as indicators for whether the student was enrolled in an honors level math class or in a geometry class. With these variables, there is a statistically significant relationship between the classroom environment index and math GPA with students who report higher levels of support and challenge from teachers in their math classes also receiving higher grades. Further, students in an honors math class receive grades that are, on average, about 0.42 grade points higher than students not in honors classes. However, these reports and indicators of classroom experiences do not change the magnitude of the estimated relationship between gender and math GPA (the estimated male coefficient is -0.24 grade points).

In Model 5, we consider all the aforementioned factors together. However, in this model, the male coefficient remains almost identical to the unconditional model (-0.25 grade points). This suggests that even after holding constant observed academic performance, behaviors and habits, and classroom experience factors, and comparing students from the same cohort at the same high school, girls still outperform boys in ninth-grade math classes by about 0.25 grade points. Across all models, the relationship between gender and math GPA remains almost unchanged, suggesting that none of the student factors that we account for can explain the GPA gap. In other words, even among girls and boys who report similar levels of incoming achievement, behaviors and habits, and experiences in their math classes, there is still a sizable gender gap in grades. This finding is surprising given the observed gender differences in some of these factors and because these factors themselves are strongly related to GPA.

While this finding holds at the average of the data, it may be the case that the gender gap may be different for students with lower or higher GPAs. Particularly because boys are more likely to be represented at the top of the math test score distribution (Baye & Monsieur, 2016; Fan, Chen & Matsumoto, 1997), understanding how the relationship between gender and GPA may vary at different points in the GPA distribution is important. To explore this more fully, we used a quantile regression approach. We specified the model (as shown in Model 5 in Table 3), controlling for academic performance, behaviors and habits, and classroom experiences, as well as school-by-year fixed effects. Table 4 shows the results from the quantile regression. At the 10th-, 25th-, 50th, and 75th-percentiles, the estimated gender gap in grades are not statistically different from each other. The point estimates of the male coefficient range from -0.22 to -0.28 at the 75th- and 25th-percentiles respectively. However, for students at the top of the GPA distribution, the gap is much smaller. While girls still earn higher math grades, on average, than boys at the 90th-percentile of the math GPA distribution, the estimated coefficient on male is only -0.13, suggesting that the gender gap in math grades is smaller among top-scoring students than at other places in the grade distribution.

C. Classroom-level factors associated with smaller gender differences in math grades

We find that in the average math classroom, girls earn grades that are 0.25 grade points higher than boys. After considering multiple student-level academic and behavioral measures, we now turn to exploring aspects about math classes that might help to explain that difference. Specifically, we consider classroom composition, type and level of the math course, and teacher characteristics. Table 5 shows the results. We add these three groups of variables individually and then all together in Model 4.

First, Model 1 looks at the relationship between classroom composition and the gender gap in grades. In classes that are higher achieving, the gender gap is smaller. That is, an increase in average test scores at the class level of 1 standard deviation is associated with a 0.07-grade point decrease in the gender gap. In Model 2, we consider the relationship between course type and level and the gender gap in math grades. In honors courses, the classroom-level gender gap is 0.11 grade points smaller. Similarly, the gap is smaller in geometry classes by 0.13 grade points relative to algebra classes. Model 3 explores teacher characteristics. We find that in math classes taught by male teachers the math GPA gap is 0.10 grade points smaller than in classes taught by female teachers. We do not find statistically significant relationships between teachers' race/ethnicity or years of experience and the gender gap. (Appendix Table 4 shows how teacher gender and race/ethnicity are distributed across different levels and types of math courses.)

When included in the model simultaneously, as shown in Model 4, the findings generally remain the same. Taken together, these results suggest that classrooms with more advanced students - both in terms of incoming achievement and the level and type of the class - tend to have smaller gender gaps in grades. This is consistent with the quantile regression results showing that the gender gap is smaller for students at the top of the math GPA distribution. Additionally, this analysis provides evidence that classes taught by male teachers have smaller gender gaps in grades, so perhaps male teachers are better able to make connections with boys or are more in tune with the academic and behavioral needs of boys. We discuss implications of these findings and directions for future research in the following section.

V. Discussion

In Chicago, girls outperform boys in ninth-grade math by more than one-quarter of a grade point. Academic performance in ninth grade is a strong predictor for whether a student graduates from a high school, a key educational milestone. Course failures in ninth grade can mean retention of students and repeating material, putting students off track to persist in their education. Further, ninth-grade GPA is highly correlated with GPA in later years (Authors, 2017), which is becoming more and more critical in college applications, particularly with the movement to rely less on standardized test score measures of academic achievement. Taken together, the gender gap in GPA has real implications in terms of educational attainment for boys who may need more support to succeed in high school.

In this paper, we set out to understand more about the reasons that boys have meaningfully lower ninth-grade math grades than girls, especially given that their prior performance on math tests was similar on average. With rich data on various aspects of prior academic performance, behaviors and habits, and student reports on educational experiences in their math classrooms, we have direct measures for numerous factors that teachers often point to as explanations for the gender gap in grades. Many of these factors –specifically incoming achievement– can typically explain differences in educational outcomes that we observe in education research. In fact, we found that almost all the characteristics and measures we can account for in the data were strong predictors of math GPA; for example, students with reports of stronger academic effort earned higher grades than students with weaker study habits, and students who experienced suspensions received lower grades than students without suspensions. Further, we also documented large gender differences in these factors with girls, on average, displaying stronger academic habits and fewer behavioral problems. These findings are consistent with prior literature that show boys as more disengaged with school, both in terms of their school behavior and higher rates of disciplinary issues (DiPrete & Buchmann, 2013; Finn & Servoss, 2014; Skiba et al., 2002), as well as their weaker study habits and work completion (DiPrete & Buchman, 2013; Gershenson & Holt, 2015; Lam et al., 2012).

Based on these results, we expected that gender differences on these factors would go a long way in explaining the gender gap in grades; however, none of these academic, behavioral, or classroom experience factors were able to explain the significant, large, and persistent difference in math grades between girls and boys. These results were surprising since prior research has shown that girls' greater effort and self-discipline in schoolwork is strongly related to their favorable academic outcomes in comparison to boys (DiPrete & Jennings, 2012; Downey & Von Vaugt, 2005; Duckworth & Seligman, 2006; Lam et al., 2012). While our study used student reports on academic behaviors and classroom experiences, some studies (e.g. DiPrete & Jennings, 2012; Downey & Von Vaugt, 2005) rely on teacher's ratings instead. It may be the case that teacher perceptions of students' behaviors differ significantly from students' own views and the latter is not accurately captured on teacher-assigned grades. Including student and teachers reports along with other non-survey measures of engagement and study habits will increase the validity of these measures in future studies.

However, when we explore classroom-level factors more deeply, we are able to get a better understanding of the contexts that are more conducive to the academic success of boys. Specifically, we find that the gender gap is smaller in more advanced coursework like honors math classes or geometry math classes. That's consistent with the finding that the gender gap is smaller at the top of the grades distribution. That raises questions and points to the need for qualitative research to understand what about the classroom environment in these types of classes are more conducive to the success of boys. Perhaps, teacher beliefs and expectations may differ across levels of math classes. Riegle-Crumb and Humphries (2012) found that in advanced classes teachers were less likely to judge the class as "too difficult" for their female students. It could be the case that prerequisites for enrolling in this type of classes works as an indication of the high-level skills that students bring with them, hence gender stereotypes are not as salient. Another plausible explanation is that girls and boys respond differently to challenging instructional contexts.

Further, the finding that the gender gap is smaller in classes taught by male teachers is also worth more exploration. Do male teachers have more developmentally appropriate expectations of boys than do female teachers? Are they better able to relate to the boys in their classrooms? Prior research provides some interesting insights into the teacher-student gender patterns. For instance, Mullola et al. (2012) found that male teachers were more likely to perceive boys more positively and more capable compared to female teachers. Also, it may be that students feel more comfortable when having teachers of their own sex. Dee's research (2007) found that girls were less likely to report being disengaged in class when the teacher was female. Perhaps there is a benefit to boys of having a male role model in the classroom.

Another area for further research is in understanding better teachers' grading practices. Our findings in this evidence point to the need to continue to understand what about the nature of high schools, or even education systems more broadly, results in poorer outcomes for boys than for girls. Particularly, researchers, policymakers, and practitioners should look carefully into behaviors inside classrooms to understand how teacher grading practices (i.e., how teachers make decisions about their assignments and what to weigh more heavily in determining final grades), classroom dynamics (e.g. interactions between students and teachers), and school structures may be contributing to the gender gap in GPA. These are empirical questions that warrant further exploration, which will provide insights to district and school leaders as well as teachers about the grading policies and practices that would result in more equitable grading for boys.

Notes

1. Our analysis focused on students attending traditional CPS high schools. Charter schools are not required to report data on grades to the district, hence we excluded from our sample students who enrolled in this type of school. Additionally, non-traditional high schools (formerly known as "alternative") often follow different school structures and target very high-risk student populations, making those schools sufficiently different to warrant exclusion from the analysis.
2. This slight gender imbalance occurs because girls are somewhat more likely to respond to the survey (87% of first-time ninth-grade girls with math grades answered the survey, compared to 83% of boys).
3. Letter grades were converted to numeric grades as follows: A=4, B=3, C=2, D=1, and F=0. Across two academic semesters, the majority of students in our analytic sample (87%) took two math courses (one per semester). An additional 12 percent of the sample took up to four math courses (two courses each semester). Only 3 students took more than four math courses during the academic year.

4. In our sample, NWEA MAP Math and Reading scores were strongly correlated (r=0.74). Hence, we decided to average the two scores in a single achievement measure to proxy students' incoming skills when starting ninth grade.
5. Most students who are missing 8th-grade test scores entered CPS in 9th grade but were not active CPS students in 8th grade. For students with 8th-grade test scores, we regressed 8th-grade math test scores on all observed ninth-grade covariates (i.e., student demographics, behaviors, survey reports). For the subsample of students who were missing test scores, we generated a predicted 8th-grade test score by applying the coefficient estimates from the regression to the covariate values for each student. In our main regression models, we use this predicted value for the students who are missing the test score, and we also include an indicator equal to one for students with missing test scores and 0 otherwise.
6. In the spring of 2016-17 and 2017-18 school years, 85 percent of first-time ninth graders with math grades responded to the 5Essentials survey. Details about the 5Essentials survey and the methodology for constructing survey measures can be found at: http://ccsr.uchicago.edu/surveys.
7. Among survey responders, some students did not respond to all survey measures. Missing rates ranged from 0.5% to 1.6%. However, *Higher-order Math Instruction* survey measure had a higher rate of missingness (8%) likely due to student attrition (i.e. items in the 2017 survey were positioned at the end of the questionnaire). We imputed missing survey responses using a regression-based approach. We regressed each survey measure on all observed ninth-grade covariates and used those predicted survey scores to replace

missing values. We also included in our models a missing data indicator that equals one for students who were missing survey measures and had imputed values.

8. Math classroom measures were highly correlated (r = 0.56 to 0.69). Our Principal Component Analysis (PCA) supported a one-component structure that grouped all measures into one index (72% of variance explained with factor loadings ranging from .479 to .516). We used the imputed version of the survey measure to create the Class Environment Index as the simple average of these measures.
9. Because most students in CPS qualify for free/reduced price lunch, using an SES measure tied to a student’s residential Census block provides more variation in student economic conditions. We note that it is just a proxy for an individual's own circumstances and is a more accurate measure in parts of the city where there is less heterogeneity in SES within a Census block.
10. Among ninth graders taking math classes in CPS we do find that girls start their freshman year with a stronger math preparation (0.05 SD average difference in 8th grade math test scores). However, in our study sample, we have boys with higher incoming math skills compared to the overall CPS population.

References

Authors (2017).

Authors (2020).

Allensworth, E. M., & Clark, K. (2020). High School GPAs and ACT Scores as Predictors of College Completion: Examining Assumptions About Consistency Across High Schools. *Educational Researcher*, *49*(3), 198-211. https://doi.org/10.3102/0013189X20902110

Allensworth, E.M., Healey, K., Gwynne, J.A., & Crespin, R. (2016). High school graduation rates through two decades of district change: The influence of policies, data records, and demographic shifts. Chicago, IL: University of Chicago Consortium on School Research.

Baye, A. & Monsieur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-scale Assessments in Education, 4*(1), 2-16. https://doi.org/10.1186/s40536-015-0015-x

Bowen, W., Chingos, M., & McPherson, M. (2009). *Crossing the Finish Line: Completing College at America's Public Universities*. PRINCETON; OXFORD: Princeton University Press.

Bowers, A.J. (2011) What's in a Grade? The Multidimensional Nature of What Teacher Assigned Grades Assess in High School. *Educational Research & Evaluation, 17*(3), 141-159. https://doi.org/10.1080/13803611.2011.597112

Brookhart, S.M., Guskey, T.R., Bowers, A.J., McMillan, J.H., Smith, J.K., Smith, L.F., Stevens, M.T., & Welsh, M.E. (2016). A century of grading research: meaning and value in the most common educational measure. *Review of Educational Research, 86*(4), 803-848. https://doi.org/10.3102/0034654316672069

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *The Journal of Human Resources*, *42*(3), 528–554. https://doi.org/10.3368/jhr.XLII.3.528

DiPrete, T.A., & Buchmann, C. (2013). *The Rise of Women: The Growing Gender Gap in Education and What it Means for American Schools.* New York NY: Russell Sage Foundation.

DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, *41*(1), 1–15. https://doi.org/10.1016/j.ssresearch.2011.09.001

Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *The Sociological Quarterly*, *46*(2), 299–321. https://doi.org/10.1111/j.1533-8525.2005.00014.x

Duckworth, A. L. & Seligman, M. E. P. (2006). Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores. *Journal of Educational Psychology*, *98*(1), 198 –208. https://doi.org/10.1037/0022-0663.98.1.198

Duckworth, A. Shulman, E. P., Mastronarde, A. J., Patrick, S. D., Zhang, J. & Druckman, J. (2015). Will Not Want: Self-Control Rather than Motivation Explains the Female Advantage in Report Card Grades. *Learn Indiv Differ*, *39*, 13-23. https://doi.org/10.1016/j.lindif.2015.02.006.

Duckworth, A.L, & Quinn, P.D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*, 166-174. https://doi.org/10.1080/00223890802634290

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *9*, 1087-1101. https://doi.org/10.1037/0022-3514.92.6.1087

Fahle, E. M., & Reardon, S. F. (2018). Education. In Stanford Center on Poverty and Inequality (Ed.), State of the Union: The Poverty and Inequality Report, Special Issue (pp. 9–12). Pathways Magazine.

Fan, X., Chen, M. & Matsumoto, A. R. (1997). Gender differences in mathematics achievement: Findings from the National Education Longitudinal Study of 1988. *The Journal of Experimental Education, 65*(3), 229-242. https://doi.org/10.1080/00220973.1997.9943456

Farrington, C.A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T.S., Johnson, D.W., & Beechum, N.O. (2012). *Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago, IL: University of Chicago Consortium on Chicago School Research.

Finn, J. D. & Servos, T. J. (2014). Misbehavior, Suspensions and Security Measures in High School: Racial/ethnic and gender differences. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5* (2-11), 1-50. Retrieved from: https://digitalcommons.library.tmc.edu/childrenatrisk/vol5/iss2/11

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109. https://doi.org/10.3102/00346543074001059

Gershenson, S. & Holt, S. (2015). Gender gaps in high school students' homework time. *Educational Researcher*, *44*(8), 432–441. https://doi.org/10.3102/0013189X15616123

Gherasim, L. R., Butnaru, S. & Mairean, C. (2012): Classroom environment, achievement goals and math's performance: gender differences. *Educational Studies*, 1-12. https://doi.org/10.1080/03055698.2012.663480

Gibb, S. J., Fergusson, D. M & Horwood, L.J. (2008). Gender differences in educational achievement to age 25. *Australian Journal of Education, 52*(1), 63-80. https://doi.org/10.1177/000494410805200105

Gil Gomes Carvalho, R. (2016). Gender Differences in academic achievement: The mediating role of personality. *Personality and Individual Differences*, *94*, 54-58. https://doi.org/10.1016/j.paid.2016.01.011

Green, J. P. & Winters, M. A. (2006). Leaving Boys Behind: Public High School Graduation Rates. *Civic Report No. 48*. Manhattan Inst., New York, NY. Center for Civic Innovation. Retrieved from: https://files.eric.ed.gov/fulltext/ED491633.pdf

Goldin, C., Katz, L. F. & Kuziemko, I. (2006). The Homecoming of American College Women: The Reversal of the College Gender Gap. *Journal of Economic Perspectives*, *20*(4), 133–156. https://doi.org/10.1257/jep.20.4.133

William and Flora Hewlett Foundation. (2013). *Deeper learning competencies*. Retrieved from: http://www.hewlett.org/library/hewlett-foundation-publication/ deeper-learning-defined

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*, 625-638. https://doi.org/10.1111/1467-8624.00301

Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., & Michelman, V. (2017). The struggle to pass algebra: Online vs. face-to-face credit recovery for at-risk urban students. *Journal of Research on Educational Effectiveness*, *10*(2), 272-296. https://doi.org/10.1080/19345747.2016.1168500

Hyde, J.S., Lindberg, S.M., Linn, M.C., Ellis, A.B. & Williams, C.C. (2008). Gender similarities characterize math performance. *Science, 321*, 494 – 495. https://doi.org/10.1126/science.1160364

Jacob, B. (2002). Where the Boys Aren't: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education. *Economics of Education Review, 21*, 589-598. https://doi.org/10.1016/S0272-7757(01)00051-6

Klem, A. M. & Connell, J. P. (2004). Relationships Matter: Linking Teacher Support to Student Engagement and Achievement. *Journal of School Health*, *74*(7), 262-273. https://doi.org/10.1111/j.1746-1561.2004.tb08283.x

Lam, S., Jimerson, S., Kikas, E., Cefai, C., Veiga, F. H., Nelson, B., . . . Zollneritsch, J. (2012). Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *Journal of School Psychology, 50*, 77-94. https://doi.org/10.1016/j.jsp.2011.07.004

Lietaert, S., Roorda, D., Laevers, F., Verschueren, K. & De Fraine, B. (2015). The gender gap in student engagement: The role of teachers' autonomy support, structure, and involvement.

British Journal of Educational Psychology, *85*(4), 498-518. https://doi.org/10.1111/bjep.12095

Lindberg, S.M., Hyde, J.S., Petersen, J.L. & Linn, M.C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychological Bulletin*, *136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Middleton, M. J., & Midgley, C. (2002). Beyond motivation: Middle school students' perceptions of press for understanding in math. *Contemporary Educational Psychology, 27*(3), 373–391. https://doi.org/10.1006/ceps.2001.1101

Mullola, S., Ravaja, N., Lipsanen, J., Alatupa, S., Hintsanen, M., Jokela, M. & Keltikangas-Järvinen, L. (2012). Gender differences in teachers' perceptions of students' temperament, educational competence, and teachability. *British Journal of Educational Psychology*, *82*(2),185-206. https://doi.org/10.1111/j.2044-8279.2010.02017.x.

O'Dea, R.E., Lagisz1, M., Jennions, M.D. & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to t expected patterns for STEM. *Nature communications*, *9*, 1-8. https://doi.org/10.1038/s41467-018-06292-0

Perkins, R., Kleiner, B., Roey, S., & Brown, J. (2004). *The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990-2000. NCES 2004-455*. National Center for Education Statistics. New York, NY: Diane Publishing.

Reardon, S.F., Fahle, E.M., Kalogrides, D., Podolsky, A., & Zarate, R.C. (2018). *Gender Achievement Gaps in U.S. School Districts* (CEPA Working Paper No.18-13). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp18-13.

Riegle-Crumb, C. & Humphries, M. (2012). Exploring Bias in Math Teachers' Perceptions of Students' Ability by Gender and Race/Ethnicity. *Gender & Society*, *26*(2), 290-322. https://doi.org/10.1177/0891243211434614

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, *81*(4), 493–529. https://doi.org/10.3102/0034654311421793

Samuelsson, M. & Samuelsson, J. (2016). Gender differences in boys' and girls' perception of teaching and learning mathematics. *Open Review of Educational Research*, *3*(1), 18-34. https://doi.org/10.1080/23265507.2015.1127770

Skiba, R., Michael, R. Nardo, A. C. & Peterson, R. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review, 34*(4), 317–342. https://doi.org/10.1023/A:1021320817372

Stetser, M., and Stillwell, R. (2014). *Public High School Four-Year On-Time Graduation Rates and Event Dropout Rates: School Years 2010–11 and 2011–12. First Look* (NCES 2014-391). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

Tennant, J. E., Demaray, M. K., Malecki, C. K., Terry, M. N., Clary, M., & Elzinga, N. (2014). Students' Ratings of Teacher Support and Academic and Social–Emotional Well-Being. *School Psychology Quarterly*, 1-21. https://doi.org/10.1037/spq0000106

U.S. Department of Education, National Center for Education Statistics. (2017). Table 227.40: Percentage of elementary and secondary school students who do homework, average time spent doing homework, percentage whose parents check that homework is done, and percentage whose parents help with homework, by frequency and selected characteristics: 2007, 2012, and 2016. In U.S. Department of Education, National Center for Education Statistics (Ed.), *Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (PFI-NHES:2007, 2012, and 2016).* Retrieved from https://nces.ed.gov/programs/digest/d17/tables/dt17_227.40.asp

Voyer, D. & Voyer, S.D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140* (4), 1174-1204. https://doi.org/10.1037/a0036620

Warren, J. R., Hoffman, E., & Andrew, M. (2014). Patterns and Trends in Grade Retention Rates in the United States, 1995–2010. *Educational Researcher*, *43*(9), 433–443. https://doi.org/10.3102/0013189X14563599

Wilcox, G., McQuay, J., Blackstaffe, A., Perry, R. & Hawe, P. (2018). Supporting Academic Engagement in Boys and Girls. *Canadian Journal of School Psychology, 33*(3), 179-192. https://doi.org/10.1177/0829573517703239

Tables & Figures

Table 1. Student characteristics

	CPS First-time Ninth Graders (N=37,404)	Analytic Sample (N=9,585)
Female	51%	51%
Black	32%	28%
Latinx	49%	50%
White	11%	12%
Other	8%	9%
Low-SES	33%	29%
High-SES	34%	37%
Has an Individualized Education Plan	14%	12%
English Learner	9%	10%
Homeless	4%	2%
Math test score in 8th grade (s.d.)	238.1 (18.5)	239.7 (18.4)
Reading test score in 8th grade (s.d)	226.5 (14.1)	227.2 (14.0)
Math GPA in 9th grade (s.d.)	2.53 (1.11)	2.64 (1.08)

Notes. Student socioeconomic status (SES) is measured at the Census block group level based on the student's residential address. Block-group SES is calculated as an index of the percentage of adult males employed and the percentage of families with incomes above the poverty line in the Census block group. All 8th-grade students in CPS take the NWEA in the spring of each academic year, and the scores shown in this table are the raw scale scores.

Table 2. Classroom characteristics

	Math Classrooms (N=1,281)
More than 50 percent of students are male	43%
Average 8th-grade test scores in z-scores (s.d.)	-0.04 (0.91)
Type of math class:	
Algebra	81%
Geometry	14%
Advanced	4%
Electives	1%
Honors class	40%
Taught by female teacher	44%
Taught by white teacher	53%
Taught by Black teacher	14%
Taught by Latinx teacher	18%
Average teachers' years of experience (s.d.)	11 (7.4)

Notes. The typical course for ninth graders in CPS is algebra; students may be enrolled in a higher-level math course if they completed algebra prior to entering high school. Teacher gender and race/ethnicity is missing in 2 percent of classrooms. Teacher experience is measured as years teaching in CPS.

Table 3. Student-level factors associated with math grades

Outcome: Ninth-grade math GPA	(1) Gender	(2) Academic Preparation	(3) Behaviors and Habits	(4) Classroom Experience	(5) All
Male	-0.253***	-0.284***	-0.235***	-0.242***	-0.251***
	(0.030)	(0.025)	(0.028)	(0.027)	(0.025)
8th-grade test scores in math and reading (z-score)		0.533***			0.433***
		(0.031)			(0.028)
Attendance rate (log z-score)			0.348***		0.313***
			(0.019)		(0.016)
Any out-of-school suspension			-0.340***		-0.296***
			(0.048)		(0.046)
Self-report of academic effort (z_score)			0.107***		0.102***
			(0.012)		(0.012)
Self-report of persistence (z_score)			0.128***		0.106***
			(0.011)		(0.011)
Honors level				0.422***	0.0315
				(0.054)	(0.051)
Geometry class				0.296***	0.0910
				(0.072)	(0.070)
Math instruction (z score)				-0.0296*	-0.0399**
				(0.013)	(0.012)
Class environment index (z score)				0.157***	0.0428**
				(0.016)	(0.014)
Intercept	2.761***	2.625***	2.741***	2.435***	2.636***
	(0.014)	(0.021)	(0.026)	(0.031)	(0.031)
N	9585	9585	9585	9585	9585
R^2	0.171	0.301	0.357	0.256	0.429

Notes. The outcome is GPA across all the semester-level math courses the student took in ninth grade. Models (2)-(5) include demographic controls at the student level (race/ethnicity, special education, ELL and homeless status, and Census block-level SES). With the exception of the SES variables and homeless status indicator, the other student baseline characteristics are statistically significant predictors of math GPA. The indicators for honors/ geometry indicate that at least one of the math courses a student took was an honors/geometry class. 87 percent of students in the sample took only one math course. All models include school-by-year fixed effects. Standard errors clustered at the school level shown in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.00$.

Table 4. Adjusted gender differences in Math GPA at different points of the performance distribution

	(1)	(2)	(3)	(4)	(5)
	10^{th} percentile	25^{th} percentile	50^{th} percentile	75^{th} percentile	90^{th} percentile
Male	-0.265^{***}	-0.276^{***}	-0.246^{***}	-0.218^{***}	-0.130^{***}
	(0.034)	(0.027)	(0.023)	(0.023)	(0.016)
Intercept	2.120^{***}	2.300^{***}	3.239^{***}	3.387^{***}	3.906^{***}
	(0.222)	(0.295)	(0.275)	(0.266)	(0.397)
N	9585	9585	9585	9585	9585

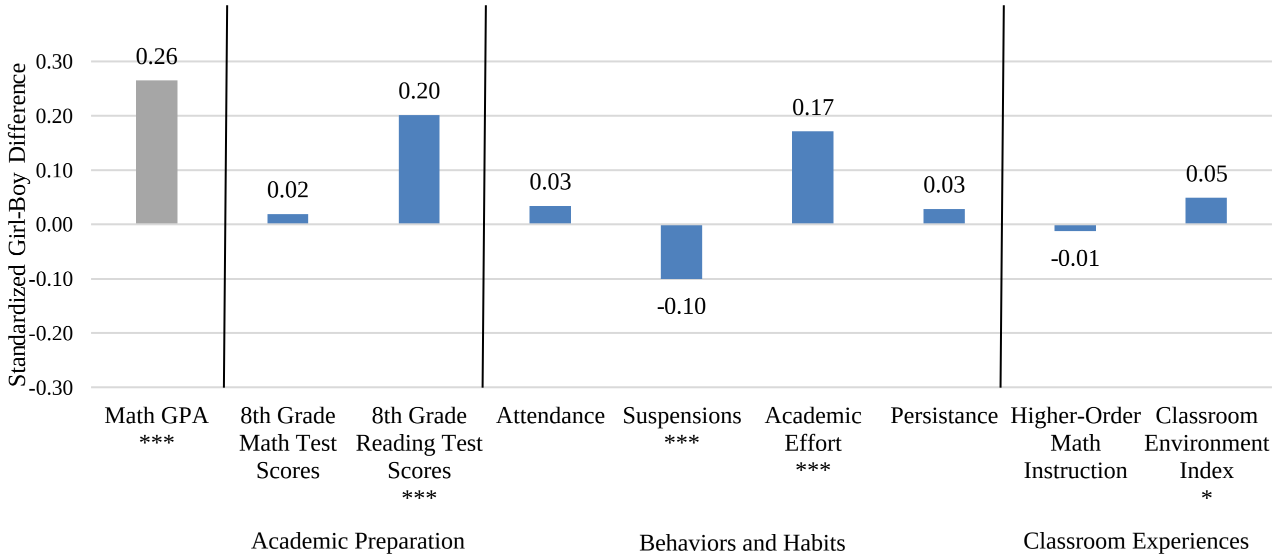
Notes. The results shown in the table come from a quantile regression where the outcome is GPA across all the semester-level math courses the student took in ninth grade. When testing for statistical differences at different points in the GPA distribution, only the male coefficient at the 90^{th}-percentile is statistically different from the male coefficients at other points in the GPA distribution. Models control for students' demographic characteristics, 8th grade average test scores, honors and geometry indicators, and all the academic preparation, behavioral, and classroom experience measures included in the full model specification shown in Table 3. All models include school-by-year fixed effects. Standard errors clustered at the school level shown in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.00$.

Table 5. Classroom-level factors associated with math grades

Outcome: classroom-level gender gap in math GPA	(1) Class composition	(2) Course type and level	(3) Teacher characteristics	(4) Full model
Proportion of students who are male	0.150			0.141
	(0.119)			(0.120)
Average 8th-grade test score (z score)	-0.072*			-0.00266
	(0.032)			(0.044)
Honors		-0.111**		-0.107*
		(0.036)		(0.045)
Type of math class				
Geometry (n=181)		-0.125**		-0.133**
		(0.047)		(0.048)
Advanced (n=47)		-0.173		-0.199
		(0.132)		(0.127)
Electives (n=13)		-0.251***		-0.196**
		(0.072)		(0.071)
Teacher characteristics				
Male			-0.102**	-0.120***
			(0.036)	(0.035)
Black			0.132	0.118
			(0.072)	(0.075)
Latinx			0.048	0.011
			(0.058)	(0.058)
Other Race/Ethnicity			0.043	0.033
			(0.067)	(0.064)
10 to 20 years of experience			-0.037	-0.038
			(0.045)	(0.046)
More than 20 years of experience			0.003	0.004
			(0.053)	(0.055)
Intercept	0.177**	0.325***	0.286***	0.304***
	(0.058)	(0.014)	(0.030)	(0.065)
N	1281	1281	1281	1281
R^2	0.077	0.084	0.081	0.100

Notes. The outcome is the girl-boy difference in the final math GPA aggregated to the classroom level. The average unadjusted gender gap in final grade at the classroom level is 0.25 grade points. Average 8th-grade test score is aggregated across students at the classroom level. Teacher gender and race/ethnicity is missing in 2 percent of classrooms. Teacher experience is measured as years teaching in CPS. Standard errors clustered at the school level shown in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 1. Standardized differences between girls and boys on ninth-grade math GPA and other factors



Note. Bars above 0.0 indicate that girls have average higher values than boys; bars below 0.0 indicate that boys have average higher values than girls. All variables were standardized within school and year to have a mean of 0 and a standard deviation of 1. Outcomes with an * denote statistically significant differences between girls and boys. The raw average ninth-grade math GPA is 2.78 for girls and 2.49 for boys.