Gender Gaps in College Enrollment: The Role of Gender Sorting Across Public High Schools

Dylan Conger and Mark C. Long

This article uses Florida administrative data to evaluate the role that public high schools play in the growing female advantage in college enrollment. We first show evidence of gender sorting across public high schools that is beyond what one would observe if students were randomly assigned to their schools. Using regression and decomposition techniques, we then find that across-school gender sorting explains 12% and 16% of females’ higher rates of enrollment among Hispanic and Black students, respectively. This relatively large contribution of high schools to gender disparities in college enrollment among Black and Hispanic students has implications for educators at all levels.

Keywords: achievement gap; educational policy; gender disparities; college enrollment; postsecondary education; regression analyses; school choice; school effects; secondary data analysis; segregation gender studies

Why Study Gender Sorting?

Launched by the famous Coleman et al. (1966) report, a large literature has evolved around the relationship between school sorting and racial or socioeconomic educational disparities (for recent examples, see Crosnoe, 2009; Hanushek, Kain, & Rivkin, 2009), but relatively little attention has been devoted to the impact of school sorting between boys and girls on disparities in their educational attainment. This hole in the recent empirical literature has a few explanations. First, modern-day gender disparities in educational outcomes tend to be smaller than racial and socioeconomic disparities, and thus gender gaps in outcomes may generate relatively less concern among educators and policymakers. Second, unlike sorting by race and class, which are often driven by discrimination or limited economic mobility, it’s not clear that sorting by gender stems from any inequalities or injustices. Third, single-sex schools are relatively uncommon in the United States, with less than 2% of students attending schools that serve only one gender (Long & Conger, 2013). The rarity of such schools provides researchers with few reasons to investigate gender sorting, especially in the public school system.

By now, most educators and researchers in highly industrialized nations are familiar with the growing female advantage in college enrollment. In fall 2010, only 43% of enrollees in U.S. 4-year and 2-year postsecondary institutions were male and the National Center for Education Statistics (NCES) projects an enrollment increase of 21% for women relative to only 12% for men through 2019 (Hussar & Bailey, 2011; NCES, 2012). These gender disparities are even larger among some racial groups; in fall 2010, for instance, 45% of White undergraduates were male in comparison to only 42% of Hispanic undergraduates and 37% of Black undergraduates (NCES, 2012). The handful of prior studies that attempt to explain females’ higher rates of enrollment typically point to their higher high school grades, rigorous course-taking, higher reported postsecondary ambitions, and higher probability of graduating from high school (Conger & Long, 2010; Jacob, 2002; Peter & Horn, 2005; Reynolds & Burge, 2008; Riegle-Crumb, 2010). Prior research has also found that girls may benefit from higher parental, peer, and teacher expectations and demonstrate stronger noncognitive skills, such as organization, self-discipline, attentiveness, dependability, and seeking help from others (DiPrete & Jennings, 2012; Jacob, 2002; Reynolds & Burge, 2008; Riegle-Crumb, 2010). Yet previous studies have been unable to entirely explain the female advantage in college-going. In this article, we take an unprecedented look at the role of sorting between boys and girls across public U.S. high schools in explaining the female advantage in 4-year college enrollment.

1Trachtenberg School of Public Policy and Public Administration, The George Washington University, Washington, D.C.
2Daniel J. Evans School of Public Affairs, University of Washington, Seattle, WA
Recent trends in education, however, suggest that some attention be paid to the gender composition of the nation's schools. To begin, amendments to Title IX provide public schools and school districts with more flexibility than in previous years to offer single-sex classrooms and schools, and recent evidence suggests that they are on the rise (Schema, 2006; U.S. Department of Education [USDOE], 2008). In addition, several recent studies document evidence of modest, but nonrandom, amounts of gender sorting within co-educational public and private schools across the nation (Corcoran & Jennings, 2011; Long & Conger, 2013). Such sorting may have implications for the education of boys and girls. Classroom observations, for instance, reveal that male students are often more vocal and distracted than female students, which can occupy teacher time and detract from the learning environment (e.g., Younger, Warrington, & Williams, 1999). The few studies that attempt to quantify the effect of male peers in classrooms and schools on student achievement also primarily find negative effects (Hoxby, 2000; Whitmore, 2005). Thus, if boys and girls attend different schools, even if the quality of the teachers and curricula are equal, the greater exposure of boys to other boys could lower their achievement and college-going.

Further, even if schools do not cater to one gender, variations in the characteristics of the schools that males and females attend could matter to their educational outcomes. Different schools produce different college-going climates due to a number of school attributes, such as the instructional ability of the teachers, the ambitions of the students, and the knowledge and organization of guidance counselors (Hill, 2008; McDonough, 1997). Correspondingly, if females are more likely than males to attend schools with highly qualified teachers or strong academic programs, such differences could widen gender gaps in achievement and attainment.

A Preview of the Study Methods and Findings

Our study examines the consequences of gender sorting on gender gaps in college enrollment in two stages. First, with administrative data on four cohorts of public school students in the state of Florida, we measure the degree of sorting between males and females across schools. We then compare the observed level of sorting to what we would observe by chance if students were distributed randomly across schools; we do this by randomly assigning students to schools 1,000 times and comparing the observed level of sorting to the distribution under random allocation. In the second stage, we use regression and Blinder-Oaxaca decomposition techniques to determine how much of the female advantage in college enrollment can be attributed to differences in the public high schools that boys and girls attend. Given national differences in gender gaps within racial/ethnic categories, we also examine the levels of gender sorting and the contribution of gender sorting to gender enrollment gaps among the major racial/ethnic groups in our sample: White, Black, and Hispanic students.

Consistent with recent trends nationwide, we first find that though gender sorting is much smaller than sorting by race, boys and girls in Florida sort into different public high schools at a level beyond what one would expect if students were randomly assigned to their schools. We then document that approximately 12% and 16% of gender gaps in college enrollment among Hispanic and Black students (respectively) can be linked to differences in the high schools that boys and girls attend. Among White students, gender sorting plays a much smaller role in the gender gap, with only 5% of the female advantage in college enrollment linked to their high schools. We discuss the implications of these findings for educators, policymakers, and researchers in our concluding section.

Recent Evidence on Gender Sorting Across U.S. Schools

Our contribution to this relatively new line of research on higher rates of female college-going is to ask whether the high schools that boys and girls attend matter to their differing trajectories. In order for schools to matter to gender disparities, there has to be some level of gender sorting across schools. A priori, it seems as if boys and girls in the U.S. should more or less be enrolled in the same high schools. Public school attendance is largely based on the student's residential location. Assuming that parents' residential location choices are not based on the sex of their children, then we should expect that each residential location would have roughly the same share of male and female teenagers, with any deviation occurring by random chance.

Yet there are several reasons to expect high schools to vary in their male share. First, boys and girls may, indeed, choose different schools. Surveys of parents choosing schools indicate that they consider a number of characteristics, such as student achievement, proximity to home, teacher quality, school safety, extracurricular activities, athletics programs, and the composition of the student body (Rose, 2001). In addition, we know that parent characteristics (such as education and race/ethnicity) affect the kind of information that parents have access to and how they use the information to make their choices (Henig, 1994; Piché & Taylor, 2004; Robenstine, 2001; Schneider, Teske, & Marschall, 2000). Although most of the school choice literature has focused on across-family differences, there is some evidence that parents also evaluate schools differently depending upon the gender of their child (David, 1997; Jackson & Bisset, 2005) and that boys and girls of secondary school age evaluate schools differently (Hastings, Kane, & Staiger, 2006). For instance, parents of girls prefer same-sex schools, whereas parents of boys tend to consider other school resources, such as the facilities (Jackson & Bisset, 2005). In their study of the effects of the Charlotte-Mecklenburg school choice program, Hastings et al. (2006) also find that White females are more likely than White males to enroll in academically focused schools. In a companion paper that uses the National Household Education Surveys (NHES) a U.S. survey of parents, we also find that parents of girls are significantly more likely to report that they home school than parents of boys, and parents of girls are more likely to gather information on school performance (among those who considered attending other schools) than parents who just have boys (Long & Conger, 2013). Further, we find that a large share of parents report that they considered other schooling options and/or moved to particular neighborhoods for their child's school, which suggests that there is ample scope for parents to make gender-based choices (Long & Conger, 2013).
These student and parental preferences could result in gender sorting within the public school system in two ways. First, in districts that have open enrollment programs and several high schools (including, for instance, charter schools that disproportionately serve the needs of one gender), parents may choose different schools for their daughters and sons. Second, nearby private schools may be more or less suited for children based on gender. For instance, an all-girls private school could lead to a higher share of males in the local public school. The sorting could arise either from parents choosing different schools for their daughters and sons (within-family sorting) or from across-family sorting if families with only sons are more inclined to send their boys to private schools, whereas families with only daughters are more inclined to send their girls to the local public school. Sorting by gender could be driven by tastes for same-sex peers or by the characteristics of schools that are correlated with one gender (e.g., a science and tech school may draw a disproportionate share of males). In addition, given the higher levels of school choice among Black and Hispanic parents, gender sorting could be larger among the Black and Hispanic than White and Asian populations (Grady & Bielick, 2010).

Gender sorting in the later high school grades could be further influenced by variations across high schools in male-to-female dropout rates. Even if the gender ratio is balanced in the ninth grade, schools with higher rates of male dropouts will produce higher female shares in the later grades, producing gender imbalance across schools. Dropout rates could, of course, be driven by differences in the qualities of the males (relative to females) that enroll in a particular school or to differences in the schools’ ability to graduate their male students.

Turning to the direct empirical evidence, an increasing number of studies document evidence of nonrandom amounts of gender sorting in U.S. public and private schools. Using both local and national data sources, for example, several studies find that girls disproportionately enroll in charter schools and that the disparities are larger in secondary school than in earlier grades (e.g., Abdulkadiroglu et al., 2009; Booker, Gill, Zimmer, & Sass, 2009; Corcoran & Jennings, 2011; Hoxby & Murarka, 2009). Using data on public and private school enrollments across the nation, we find in our companion paper that the gender sorting occurs not only between charter and traditional public schools, but also within traditional public schools and private schools (Long & Conger, 2013). We also demonstrate that the variation in the male share across schools (the measure that they use to demonstrate gender sorting) is much larger than what would occur if students were randomly assigned to their schools, with higher levels of nonrandom gender sorting at the high school level than the elementary school level.1 For instance, the actual standard deviation of male share across U.S. schools exceeds the random distribution in all grades, and this standard deviation is more than double the expected level in each of the Grades 9–12, whereas just 18%–46% greater than expected in Grades K–8. Moreover, we find that the sorting in the high school grades is primarily within sectors (i.e., within the public school system and within the private school system) rather than across sectors.

### Data and Analysis Sample

Our analysis aims to shed some light on the possible consequences of boys and girls choosing different high schools using a unique set of data from the Florida Department of Education. For this inquiry, we focus on the 536,985 students who graduated from a Florida public high school or who earned a General Equivalency Diploma (GED) within 4 years of high school entry in any of the years 2002–03 through 2005–06.2 We refer to this analytic sample as “on-time graduates.” In a subset of our analyses, we expand the sample to include students who were observed enrolling in any grade in a Florida public high school, and refer to this sample as “All 9th—12th graders” (N = 885,922).

For each student, the data include academic records (for example, scores on the statewide Florida Comprehensive Achievement Test, FCAT) as well as socioeconomic and demographic information. The data files also include identifiers for the over 400 high schools that students attend. Enrollment in a 4-year college (including part-time and full-time enrollment) within 5 years of high school entry is based on administrative data from Florida’s public colleges combined with National Student Clearinghouse data on students enrolling in private and out-of-state colleges. We have chosen to focus our analysis on “immediate” enrollment in 4-year colleges given the high labor market returns to bachelor’s degree attainment and evidence that students who immediately start in 4-year colleges are more likely to complete bachelor’s degrees than those who delay enrollment or start in 2-year colleges (Long & Kurlaender, 2009).

This dataset has two main advantages in answering the question that we seek to answer. First, unlike the few datasets that sample students across the nation, the census of students and schools provided by the statewide administrative dataset allows us to reliably estimate the effects of attending particular high schools (i.e., to estimate high school fixed effects) for all students and for racial/ethnic subgroups. Second, our dataset contains pre–high school test scores (i.e. eighth grade FCATs), which provide a key control variable for isolating the contribution of the high school to later outcomes.

Table 1 provides descriptive statistics on our analytic sample of on-time graduates. The first row provides 4-year college enrollment rates among males and females and reveals a female advantage of 6.7 percentage points. This disparity is very close to the gender gap in 4-year college enrollment obtained from the Education Longitudinal Study (ELS), where female high school graduates are 7.3 percentage points more likely to enroll than males (author’s calculations). The second row in Table 1 shows that the gender gap in postsecondary enrollment (at either a 2-year or 4-year institution) is even larger, 10.6 percentage points. The table also shows gender differences in eighth grade achievement that are consistent with prior research: males earn higher math scores and lower reading scores than females (Fryer & Levitt, 2010; Robinson & Lubinski, 2011). Male high school graduates are also more likely to be White and much more likely to be designated with a special need than females. Males and females have modest differences in age and recorded eligibility for free or reduced-price lunch, and no significant difference in English Language Learner (ELL) status.
Methodology

**Measuring Gender Sorting Across Schools**

We measure the unevenness in the distribution of boys and girls across schools with the standard deviation in schools’ male share of enrollment. As with all measures of unevenness, we note that even if students were randomly assigned to their schools, the expected value of our measure is likely to be nonzero due to the natural features of randomness. That is, only a deliberate effort to achieve perfect gender balance would result in evenly distributed schools, while a random allocation process is likely to deviate slightly from a perfectly even distribution.

To evaluate the extent to which the observed standard deviation in male share deviates from what would be expected if males and females were randomly assigned to schools, we construct a fictional dataset composed of the numbers of males and females enrolled in each school. We then randomly allocate the students to high schools, keeping the high schools’ enrollment fixed at their actual values. We repeat this simulation 1,000 times and report the standard deviation in male shares under random assignment. Finally, we compare the standard deviation in male share in the actual distribution to the standard deviation in the random distribution and use this “residual” to evaluate the statistical significance of gender sorting. This comparison constitutes a one-sided p value test of the hypothesis that the observed allocation results in a higher degree of gender sorting than the random allocation.3

**Calculating the Contribution of Sorting to Gender Gaps in College Enrollment**

Our analysis of the contribution of schools to gender gaps unfolds in two phases. First, we model college enrollment as a function of observable student characteristics and high school fixed effects with the following specification:

\[
\Pr[Y_{ij} = 1] = F(\alpha + \beta Male_i + X_i'\delta + H_{ij}'\gamma + e_{ij}), \quad (1)
\]

where \(Y\) is 4-year college enrollment for student \(i\) from high school \(j\), \(F()\) is the logistic distribution, and \(Male\) equals 1 if the student is male. \(X\) is a vector of student-level demographic and pre–high school achievement characteristics, including the eighth-grade test scores, race/ethnicity, age, eligibility for free or reduced-price lunch, ELL status, and special needs. \(H\) is a vector of school indicators.4 This model provides estimates of the magnitude of the relationship between each input and 4-year college-going, while holding all others constant. We pay particular attention to the coefficient on the male indicator, which indicates the gender gap in college-going, holding constant background characteristics and high schools.

For students who were enrolled in multiple high schools, we identify the high school that was attended for the most terms and set this high school’s indicator to one and all other high schools to zero. For the main analysis, we only consider enrollments in regular high schools in order to minimize the possibility that our results are driven by boys’ disproportionate attendance in irregular schools, including vocational, special education, and other alternative schools (e.g., schools in juvenile detention centers). We test for the sensitivity of these and other model choices in our robustness table.

The eighth grade test scores are missing for 25% of students in our main analysis sample, and age is missing for less than 1% of students. To address missing values for these covariates, we use multiple imputation by chained equations creating five imputed datasets. Results from the five imputed datasets are averaged, and the standard errors corrected to reflect the degree of uncertainty arising from imputing missing covariates (Royston, 2006; Rubin, 1987).

To examine the statistical significance of the contribution of high schools to the gender gap in college enrollment, we first test for the joint significance of the \(\gamma\) coefficients from Equation 1.
We then conduct a Blinder (1973)–Oaxaca (1973) decomposition to evaluate the share of the gender gap in college enrollment that can be linked to the high schools that boys and girls attend. This decomposition technique allows us to parcel the gender disparity in college enrollment into the portion that can be attributed to the background characteristics and pre–high school achievement variables that are included in Equation 1 versus the high school fixed effects that are included in Equation 1. The traditional Blinder-Oaxaca decomposition for a linear function uses the following steps. First, estimate the linear version of Equation 1 separately for boys and girls (dropping the Male indicator). Second, construct the following:

\[
(Y_f - Y_m) = (X_f - X_m)\delta_m + (H_f - H_m)\gamma_m + \hat{X}_f (\hat{\delta}_f - \hat{\delta}_m) + \hat{H}_f (\hat{\gamma}_f - \hat{\gamma}_m) + (\hat{\alpha}_f - \hat{\alpha}_m),
\]

where bars over the variables indicate sample means of the outcomes and covariates. Thus, \((Y_f - Y_m)\) is the gender gap in the mean of the college enrollment (with the \(f\) and \(m\) subscripts denoting female and male, respectively), \((X_f - X_m)\) is the vector of gender gaps of student-level characteristics, and \((H_f - H_m)\) is the vector of gender gaps of the high school indicators. \(\delta_m\), \(\gamma_m\), \(\delta_f\), and \(\gamma_f\) are the estimated coefficients from the separate male and female regressions. The key term in Equation 2 is \((H_f - H_m)\gamma_m\), which measures the contribution of differences in high school attended to gender gaps in college enrollment (assuming the enrollment returns to high school attended estimated from the male sample, \(\gamma_m\)). A simple algebraic rearrangement of Equation 2 can yield \((H_f - H_m)\gamma_f\) as the contribution of differences in high school attended to explaining gender gaps in the college enrollment (assuming female returns to high school attended, \(\gamma_f\)). As these two terms yield different results (to the extent that returns to the high school indicators vary by gender), a more agnostic approach is to assume that \((H_f - H_m)\gamma_{pooled}\) is the contribution of high schools, where \(\gamma_{pooled}\) is the vector of coefficients on the high school indicators for Equation 1 using the pooled sample of boys and girls. We use this latter approach in this article. Given the binary nature of our dependent variable, we use the analog of the Blinder-Oaxaca decomposition developed by Fairlie (2005) and the corresponding Stata command ("fairlie") by Jann (2006).\(^5\)

**Results**

We begin with a simple description of the degree to which boys and girls sort across high schools in Florida. A standard metric for measuring segregation is the Dissimilarity Index, which measures the proportion of Group A students that would have to transfer from a school in which they were overrepresented to one in which they were underrepresented to achieve an even distribution with Group B students across all schools. For the set of students observed in the 9th through 12th grades, the male-female dissimilarity index is 5%. That is, 5% of boys (or girls) would need to change schools to achieve gender balance across all schools. As expected, this degree of across-school sorting between males and females is far lower than the degree of across-school sorting between racial groups: using the same sample of students, we calculate a Black-White dissimilarity index of 48% and a Hispanic-White dissimilarity index of 54%.

We next evaluate the degree to which the sorting we observe by gender deviates from what we would expect if students were randomly assigned to their schools. For this analysis, we use the standard deviation in male share of a high school’s enrollment across schools as our sorting measure. Table 2 provides our analysis of gender sorting for the most recent class of students, the class of 2005–2006 (estimates for earlier cohorts are nearly the same and can be obtained from the authors). We examine the level of gender sorting separately for students by race/ethnicity and by whether they simply enrolled in high school and would have been on track to graduate in 2005–2006 given normal grade progression (“all 9th through 12th graders”) versus graduated from high school or with a GED within 4 years of entry (“on-time graduates”).

A number of findings are provided in Table 2. First, the amount of gender sorting is statistically different than what would occur under a random allocation of boys and girls to their schools. Among all 9th through 12th graders, for instance, the standard deviation in the male share across schools is 3.4%; if students had been randomly assigned to their schools, the standard deviation would have been 2.3%, resulting in a residual sorting of 1.1 percentage points (i.e., 46% higher than expected).\(^6\) Second, the residual level of gender sorting in Florida high schools is similar to the residual levels of sorting nationwide (using 2007–08 Common Core of Data Public Elementary/ Secondary School Universe Survey, we document a residual sorting in public high school grades that is 48-61% higher than expected; Long & Conger, 2013). Third, gender sorting is slightly higher for minority (particularly Black) students than it is for White students. For instance, among all 9th through 12th graders, the residual sorting for Black students is 1.4 percentage points compared to 0.9 percentage points for Whites. Fourth, the amount of residual gender sorting for all high school students is roughly the same as the amount of residual gender sorting for graduating students. Consistent with findings from the national studies (Corcoran & Jennings, 2011; Long & Conger, 2013), this finding suggests that a large portion of the sorting is driven by gendered choices upon enrollment as opposed to differential male dropouts across schools.\(^7\)

Table 3 provides our estimate of the contribution of the schools attended by graduating males and females to their differing rates of enrollment. The table contains selected results from our estimation of Equation 1 for the major racial/ethnic groups with all student-level covariates and high school fixed effects held constant.\(^8\) Panel A shows that gender gaps in college entry are similar for White, Black, and Hispanic on-time graduates of Florida’s public schools (with odds ratios ranging from 0.71 for Blacks to 0.76 for Hispanics). Our main interest is in whether the sorting across schools appears to mitigate or exacerbate these disparities. The high school fixed effects are jointly significant as noted by \(p\) values of less than 0.01, which is not terribly surprising given our large sample sizes.

What is more interesting is the great variation in the high school fixed effects, which illustrates how even a modest amount of gender sorting may substantially contribute to gender inequality in college entry. As an example, start with a student who
would have a 33.3% likelihood of attending a 4-year college in the year after high school if they attended the “median” high school, based on the estimated fixed effects (not shown in the table). If this student attended the 10th, 25th, 75th, or 90th percentile high school instead, the student’s estimated likelihood of enrollment would be 16%, 23%, 51%, or 80%, respectively. Thus, the high school attended is strongly associated with college enrollment, controlling for observed pre–high school characteristics.

To determine what percent of the gender gaps appear to be attributable to the high schools that females and males attend, we decompose each gender gap into the portion that can be explained by differences in the covariates. The portion of each gap that is attributed to the high school fixed effects is found in Panel B of Table 3. Differences in the high schools attended by males and females explain 10.9% of the female college-going advantage over males. That is, males appear to be disadvantaged by their high schools. Further, the high school fixed effects explain sizeable portions of the gender gaps in college entry for Blacks (15.8%) and Hispanics (12.2%), but explain a more modest portion of the gender gap for Whites (5.2%).

In Table 4, we test for the robustness of our models with two sensitivity analyses. Column 1 reprints the results from the original model, column 2 provides the results when we remove GED recipients from the sample to determine whether the results are primarily driven by higher rates of GED completion by boys, and column 3 provides the results when we include fixed effects based on enrollment terms in any high school (including alternative schools, etc.) as opposed to just “regular” high schools. Panel A of Table 4 reveals that despite changes in the number of students and schools employed in the analysis, the magnitude of the gender gap remains nearly the same. Panel B also shows that the portion of the gap that can be linked to the high school fixed effects is robust to these alternative specifications and samples.

### Discussion and Limitations

This study tackles an underexplored area—the role of gender sorting across public schools in females’ higher rate of 4-year college-going—and produces several findings that have relevance for educators, policymakers, and researchers.

**Discussion**

Consistent with a growing body of research that documents gender sorting across schools, we find that girls and boys in our case study state sort into different public schools at a level that is well beyond what one would expect if the sorting were random. In addition, boys are more likely to attend high schools that appear to disadvantage them—that is, schools that associate with a lower rate of college-going. The level of gender sorting and the disadvantage that males experience due to the sorting is higher among Black and Hispanic students than among White students. Given that we are unable to determine why boys and girls select into these different types of schools, we are reluctant to investigate or even speculate on which particular attributes of the schools seem to boost or lower college-going.

Nevertheless, the findings presented here have a number of implications. Although the degree of gender sorting that we observe in the public school system remains small relative to racial and socioeconomic sorting, recent changes to federal legislation regarding public school funding allow for greater flexibility within the public schools to cater to the needs of different genders and

---

**Table 2**

**Gender Sorting Across Public Schools in Florida, by Race/Ethnicity**

<table>
<thead>
<tr>
<th>Actual Distribution of Students</th>
<th>Random Allocation of Students</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3) (4) (5) = (3) – (4)</td>
<td></td>
</tr>
<tr>
<td>Enrollment (M)</td>
<td>Male Share (M)</td>
<td>Male Share (SD)</td>
</tr>
<tr>
<td>All 9th through 12th graders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All students</td>
<td>697</td>
<td>50.5%</td>
</tr>
<tr>
<td>White students</td>
<td>637</td>
<td>50.5%</td>
</tr>
<tr>
<td>Black students</td>
<td>682</td>
<td>50.1%</td>
</tr>
<tr>
<td>Hispanic students</td>
<td>851</td>
<td>50.8%</td>
</tr>
<tr>
<td>On-time graduates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All students</td>
<td>441</td>
<td>48.4%</td>
</tr>
<tr>
<td>White students</td>
<td>427</td>
<td>48.9%</td>
</tr>
<tr>
<td>Black students</td>
<td>399</td>
<td>47.2%</td>
</tr>
<tr>
<td>Hispanic students</td>
<td>518</td>
<td>48.2%</td>
</tr>
</tbody>
</table>

Note. (i) Data source is the Florida Department of Education, Class of 2005–06. (ii) Samples restricted to students in regular public schools. (iii) “Random allocation” results show the means from 1,000 simulations. (iv) Asterisks represent the one-tailed significance of the difference in standard deviation of “male share” using actual and random allocations: *p < .10. **p < .05. ***p < .01.
we can expect more gender sorting within the public school system (both across schools and classrooms within schools) simply due to the increasing presence of same-sex environments (Schemo, 2006; USDOE, 2008). Further, in our companion paper, we find that counties where a larger share of students attend private, magnet, charter, and irregular public schools have higher levels of gender sorting across schools in the county (Long & Conger, 2013). Thus, increasing opportunities for choice may facilitate higher amounts of gender sorting (just as they might increase sorting by race and class) in future years. Our results suggest that these trends may influence gender differences in educational outcomes and the growing female advantage in college enrollment, in particular. To set a baseline, school systems and policymakers may want to document and monitor the gender composition of the nation’s schools. If continued monitoring suggests an increase in gender sorting, and further research suggests harmful consequences to such sorting, then school systems may want to consider gender balance in their school assignment policies.

### Table 3
**The Role of High Schools in Explaining Gender Gaps in 4-Year College Entry, by Race/Ethnicity**

<table>
<thead>
<tr>
<th></th>
<th>All Students</th>
<th>White Students</th>
<th>Black Students</th>
<th>Hispanic Students</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Descriptive Statistics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of students</td>
<td>536,233</td>
<td>316,104</td>
<td>98,958</td>
<td>97,705</td>
</tr>
<tr>
<td>Male share of students</td>
<td>48.3%</td>
<td>49.4%</td>
<td>45.5%</td>
<td>47.7%</td>
</tr>
<tr>
<td>Number of high schools</td>
<td>436</td>
<td>425</td>
<td>405</td>
<td>375</td>
</tr>
<tr>
<td>Share that attended a 4-year college:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>28.9%</td>
<td>27.7%</td>
<td>25.4%</td>
<td>33.5%</td>
</tr>
<tr>
<td>Females</td>
<td>35.5%</td>
<td>34.5%</td>
<td>32.5%</td>
<td>39.8%</td>
</tr>
<tr>
<td>Difference</td>
<td>-6.7%</td>
<td>-6.8%</td>
<td>-7.2%</td>
<td>-6.3%</td>
</tr>
<tr>
<td>Odds ratio: Male odds / female odds</td>
<td>0.74</td>
<td>0.73</td>
<td>0.71</td>
<td>0.76</td>
</tr>
<tr>
<td><strong>Panel B: Decomposition Results Using High School Fixed Effects:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p value for test of joint significance of high school fixed effects</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Share of gender gap attributed to high school fixed effects</td>
<td>10.9%***</td>
<td>5.2%***</td>
<td>15.8%***</td>
<td>12.2%***</td>
</tr>
</tbody>
</table>

*Note.* (i) Sample includes all students who graduated from a regular Florida public high school or earned a GED within 4 years of high school entry in the classes of 2002–03, 2003–04, 2004–05, and 2005–06. (ii) The models are estimated using logit regressions controlling for student’s academic and demographic characteristics, as shown in Table 1. (iii) *p < .10. **p < .05. ***p < .01. (iv) Full results and standard errors are available upon request.

### Table 4
**Robustness Checks**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excluding GED recipients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school fixed effect based on enrollment in:</td>
<td>Regular HS</td>
<td>Regular HS</td>
<td>Any HS</td>
</tr>
<tr>
<td><strong>Panel A: Descriptive statistics:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of students</td>
<td>536,233</td>
<td>502,384</td>
<td>541,637</td>
</tr>
<tr>
<td>Male share of students</td>
<td>48.3%</td>
<td>47.7%</td>
<td>48.4%</td>
</tr>
<tr>
<td>Number of high schools</td>
<td>436</td>
<td>433</td>
<td>615</td>
</tr>
<tr>
<td>Share that attended a 4-year college:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>28.9%</td>
<td>30.7%</td>
<td>28.8%</td>
</tr>
<tr>
<td>Females</td>
<td>35.5%</td>
<td>37.1%</td>
<td>35.5%</td>
</tr>
<tr>
<td>Difference</td>
<td>-6.7%</td>
<td>-6.4%</td>
<td>-6.7%</td>
</tr>
<tr>
<td>Odds ratio: Male odds / female odds</td>
<td>0.74</td>
<td>0.75</td>
<td>0.74</td>
</tr>
<tr>
<td><strong>Panel B: Decomposition results using high school fixed effects:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p value for test of joint significance of high school fixed effects</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Share of gender gap attributed to high school fixed effects</td>
<td>10.9%***</td>
<td>11.5%***</td>
<td>11.8%***</td>
</tr>
</tbody>
</table>

*Note.* (i) Sample includes all students who graduated from a regular Florida public high school or earned a GED (except where indicated) within 4 years of high school entry in the classes of 2002–03, 2003–04, 2004–05, and 2005–06. (ii) The models are estimated using logit regressions controlling for student’s academic and demographic characteristics, as shown in Table 1. (iii) *p < .10. **p < .05. ***p < .01. (iv) Full results and standard errors are available upon request.
**Limitations and Suggestions for Future Research**

It is important to note that our data are limited to public school students in Florida and, as such, may not yield results that are externally valid to public schools across the nation or within other sectors. In a separate paper, we document much higher rates of gender sorting within the private school system, which suggests that our focus on public schools may provide an underestimate of the contribution of the high school to gender enrollment gaps (Long & Conger, 2013). In addition, we remind the reader that there are two possible explanations for the contribution of the high school fixed effects to gender disparities in college enrollment. The first explanation is that the gap is being created by the schools themselves—for instance, that girls are enrolling in high schools with a stronger college-going climate that boosts the likelihood of college enrollment of their students. The second possibility is that these estimated fixed "effects" absorb the influence of omitted student and/or family attributes. Our models control for several pre–high school characteristics, most importantly students eighth grade achievement scores. But if, for example, female students are more motivated than male students with the same eighth grade test scores (in ways we have not observed), and thus select into schools that place more emphasis on college-going, then the causal role of the school will be overestimated. We suspect that the true interpretation lies somewhere in between these two extremes, but we are unable to state with certainty whether the schools are producing gender gaps.

To that end, there is room for much more research on this trend. Future research should dig deeper into the sources and consequences of gender-based sorting. For instance, more surveys of parents and youth regarding their school preferences, how those preferences link to the gender of their child, and whether stated preferences match actual behavior would help pave the way for understanding the causes and consequences of gender sorting across schools and classrooms. In addition, further study should examine the role of increasing school choice options within school districts, both across and within sectors, on gender sorting across schools. Both survey research and an analysis of district-level choice options may also help to understand why the level and consequences of across-school gender sorting are higher for minorities than for White students. Research that aims to uncover the causal effects of schools on gender gaps and the specific attributes of schools that girls seem to benefit from would also move our understanding of this phenomenon to a deeper level. Continued research in this field is important given that the small amount of gender sorting we observe in our sample plays a non-trivial role in gender gaps in college enrollment, particularly among Black and Hispanic students.

**NOTES**

We thank Charles Clotfelter, Laura Evans, Dani Fumia, Dan Goldhaber, Jason Grissom, Eric Grodsky, Charlie Hirschman, Chris Jepsen, Teny Maghakian, Paco Martorell, Marta Tienda, Will Tyson, Joel Vargas, and participants at the annual meetings of the Association for Public Policy Analysis & Management, Association for Education Finance and Policy, and the Western Economic Association International for helpful comments. We thank those who participated in seminars at Stanford’s Graduate School of Education and the University of Washington’s Center for Studies in Demography and Ecology and Evans School of Public Affairs. We thank the Florida Department of Education for providing access to the data. Funding for the processing of the data (in relation to a separate research project) was provided by the U.S. Department of Education, Institute of Education Sciences, Grant R305B070131.

To isolate the portion of the total sorting that is due to nonrandom and nonmechanical influences, we used Monte Carlo simulations that randomly allocate students to schools and report the standard deviation in male share that arises from these simulations. We use the same approach in this paper.

Florida is divided into 67 counties, and each county has one school district (i.e., district and county boundaries are coterminous). The “lab schools” at the University of Florida, Florida State University, Florida A&M University, Florida Atlantic University, and the Florida Virtual School each have their own school districts. We have dropped students who primarily attended these schools from the analysis.

This approach has been used in several other studies. See, for example, Carrington and Troske (1997), Conger (2005), Hellerstein and Neumark (2005), and Long and Conger (2013).

Unlike most models of college enrollment, ours includes pre–high school academic qualifications as opposed to academic qualifications earned while in high school (such as college entrance exams, grades, and high school courses). The reason for our choice of this specification is that we want to identify the total influence of the high school on gender differences in college-going. Schools may be more equipped to prepare their students for college and such preparation may include offering more advanced courses and assisting in preparing students for college entrance exams. A model that holds constant high school academic qualifications would eliminate this mechanism from the high school fixed effect and restrict the remaining contribution of the high school to other mechanisms.

The analog to Equation 2 is as follows:

\[
\langle Y_i - T_m \rangle = \left\{ \frac{\sum_{i=1}^{N_f} F{\left( \alpha_m + X_i \delta_m + H_i \gamma_m \right)}}{N_f} \right\} \\
- \left\{ \frac{\sum_{i=N_f+1}^{N} F{\left( \alpha_m + X_i \delta_m + H_i \gamma_m \right)}}{N - N_f} \right\} \\
+ \left\{ \frac{\sum_{i=1}^{N_f} F{\left( \alpha_f + X_i \delta_f + H_i \gamma_f \right)}}{N_f} \right\} \\
- \left\{ \frac{\sum_{i=1}^{N_f} F{\left( \alpha_m + X_i \delta_m + H_i \gamma_m \right)}}{N_f} \right\}
\]

where \(N_f\) is the number of females, \(N\) is the total number of observations, and students are sorted so that observations 1 through \(N_f\) are female whereas the remainder are male.

Nearly all of this excess gender sorting occurs within districts. The standard deviation in the male share across districts is 0.95%; if students had been randomly assigned to their districts, the standard deviation would have been 0.87%, which is an insignificant difference from the actual distribution.

In an earlier version of this article, we evaluated the effects of gender sorting across high schools for all students on gender gaps in high school completion. We found that males appear to be disadvantaged by their high school, but the share of the gender gap in high school completion that was explained by high school attended was small (1.2%). Results are available from the authors on request. Since males are disadvantaged by their high schools at both stages (high school completion and college enrollment conditional on completion), the results in this article are not simply driven by our focus on those who complete high school.
We drop 762 students in high schools that have no variation in the dependent variable (college enrollment) from the regressions. Regression results can be obtained from the authors.

REFERENCES


**AUTHORS**

DYLAN CONGER, PhD, is an associate professor at the George Washington University Trachtenberg School of Public Policy and Public Administration; dconger@gwu.edu. Her research concerns economically disadvantaged and minority youth with a focus on education policies.

MARK C. LONG is an associate professor in the Evans School of Public Affairs and an adjunct associate professor in the Department of Economics at the University of Washington, Box 353055, Seattle, WA 98195-3055; marklong@uw.edu. His research examines the effects of public policies (particularly education policy) on economic opportunity and efficient social mobility, with emphasis on estimating the benefits and costs of those policies.

Manuscript received December 20, 2012
Revisions received March 29, 2013, and June 5, 2013
Accepted July 10, 2013