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

Dylan Conger
The George Washington University
Trachtenberg School of Public Policy and Public Administration
805 21st Street NW, MPA 601G
Washington, DC 20052
202-994-1456
dconger@gwu.edu

Mark C. Long
University of Washington
Daniel J. Evans School of Public Affairs
Box 353055
Seattle, WA 98195-3055
206-543-3787
marklong@u.washington.edu

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

Abstract

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 percent 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: Gender disparities, college enrollment, school choice, school effects, segregation
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 percent of enrollees in US 4-year and 2-year postsecondary institutions in were male and the National Center for Education Statistics projects an enrollment increase of 21 percent for women relative to only 12 percent for men through 2019 (NCES 2012; Hussar and Bailey 2011). These gender disparities are even larger among some racial groups; in fall 2010, for instance, 45 percent of white undergraduates were male in comparison to only 42 percent of Hispanic undergraduates and 37 percent 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 post-secondary ambitions, and higher probability of graduating from high school (Conger and Long 2010; Jacob 2002; Peter and Horn 2005; Reynolds and Burge 2008; Riegle-Crumb, 2010). Prior studies have also found that girls may benefit from higher parental, peer, and teacher expectations and demonstrate stronger non-cognitive skills, such as organization, self-discipline, attentiveness, dependability, and seeking help from others (DiPrete and Jennings 2012; Jacob 2002; Reynolds and Burge 2008; Riegle-Crumb 2010). Yet the prior research has been unable to entirely explain the female advantage in college-going.

This paper is, to our knowledge, the first to examine the role of sorting between boys and girls across public US high schools in explaining the female advantage in 4-year college enrollment. Launched by the famous Coleman report (1966), 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, and 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, gender disparities in educational outcomes tend to be smaller than racial and socioeconomic disparities, and thus gender gaps have caused 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 US, with less than 2% of students attending schools which serve only one gender (Long and Conger, Forthcoming). The rarity of such schools provides researchers with few reasons to investigate gender sorting, especially in the public school system.

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 (Schemo 2006; USDOE 2008). In addition, several recent studies document evidence of modest, but non-random, amounts of gender sorting within co-educational public and private schools across the nation (Long and Conger Forthcoming; Corcoran and Jennings 2011). 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, and 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 2009; 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.

Our study examines the consequences of gender sorting on gender gaps in college enrollment using a very simple approach. With administrative data on four cohorts of public school students in the state of Florida along with regression and decomposition techniques, we 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 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 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 percent 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 percent 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 US 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 US 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, extra-curricular 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 and Taylor 2004; Robenstine 2001; Schneider, Teske, and Marschall 2000). Though 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 and Bisset 2005) and that boys and girls of secondary school age evaluate schools differently (Hastings, Kane, and Staiger 2006). For instance, parents of girls prefer same-sex schools, while parents of boys tend to consider other school resources, such as the facilities (Jackson and 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. Using the National Household Education Surveys (NHES), a US survey of parents, Long and Conger (Forthcoming) 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. Further, Long and Conger (Forthcoming) 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.

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 while 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 and 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 9th 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 non-random amounts of gender sorting in US 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., Abdulkadirogu et al. 2009; Booker et al. 2009; Corcoran and Jennings 2011; Hoxby and Muraka 2009). Using data on public and private school enrollments across the nation, Long and Conger (Forthcoming) find that the gender sorting occurs not only between charter and traditional public schools, but also within traditional public schools and private schools. They 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 non-random gender sorting at the high school level than the elementary school level. For instance, the actual standard deviation of male share across US 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, while just 18%-46% greater than expected in grades K-8. Moreover, they 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 four years of high school entry in any of the years 2002-03 through 2005-06. 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 selected variables on the roughly 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 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 and 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. 8th grade FCATs), which provide a key control variable for isolating the contribution of the high school to later outcomes. At the same time, 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. Long and Conger (Forthcoming) 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.

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 table also shows gender differences in 8th grade achievement that are consistent with prior research: males earn higher math scores and lower reading scores than females (Fryer and Levitt 2010; Robinson and Lubienksi 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 are otherwise similar on age, eligibility for free or reduced-price lunch, and English Language Learner (ELL) status.

[Insert Table 1 here]
METHODOLOGY

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:

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

where $Y_i$ is 4-year college enrollment for student $i$ from high school $j$; $F()$ is the logistic distribution; and $Male_i$ equals 1 if the student is male. $X_i$ is a vector of student-level demographic and pre-high school achievement characteristics, including the 8th grade test scores, race/ethnicity, age, eligibility for free or reduced-price lunch, ELL status, and special needs. $H_{ij}$ is a vector of high school indicators.

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.iii We test for the sensitivity of these and other model choices in our robustness table.

The 8th grade test scores are missing for 25% of students in our main analysis sample, and other student characteristics are missing for 1% or fewer observations. 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 (Rubin 1987; Royston 2006).
To examine 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). This test determines whether high schools and their surrounding neighborhoods (or parental/student choice of high schools) affect college enrollment. We then conduct a Blinder (1973) - Oaxaca (1973) decomposition to evaluate the cumulative role of high schools attended in generating gender gaps in enrollment. 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:

\[
\begin{align*}
\bar{Y}_f - \bar{Y}_m &= (X_f - X_m) \delta_m + (H_f - H_m) \gamma_m + X_f (\delta_f - \delta_m) + H_f (\gamma_f - \gamma_m) + (\hat{\alpha}_f - \hat{\alpha}_m)
\end{align*}
\]

where bars over the variables indicate sample means of the outcomes and covariates. Thus, \( \bar{Y}_f - \bar{Y}_m \) is the gender gap in the mean of the outcome (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 returns to high school attended estimated from the male sample, \( \hat{\gamma}_m \)). 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 outcome (assuming female returns to high school attended, \( \hat{\gamma}_f \)). 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 $\mathbf{\hat{Y}}_{\text{pooled}}$ 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 paper. 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).iv

It is important to acknowledge that while we assume that the contribution of the high school fixed effect to gender disparities in college enrollment is primarily driven by the schools themselves, there is also the possibility that these fixed effects capture the influence of omitted student and/or family attributes. Our models control for several pre-high school characteristics, most importantly, students 8th grade achievement scores. But if, for example, female students are more motivated than male students with the same 8th grade test scores (in ways we have not observed), and thus select into schools that place more emphasis on college-going, then the role of the school will be overestimated. Future research on the unique contribution of the school to gender gaps would be enriched by designs that employ natural experiments, such as lotteries.

RESULTS
We begin with a simple description of the degree to which boys and girls sort across high schools in Florida using the standard deviation in male share of a high school’s enrollment across schools as our sorting measure. 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.v

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 2005-06 given normal grade progression ("All 9th-12th graders") versus graduated from high school or with a GED within four years of entry ("On-time graduates").

A number of findings emerge from 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-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).vi 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, Long and Conger (Forthcoming) document a residual sorting in public high school grades that is 48-61% higher than expected). Third, gender sorting is slightly higher for minority (particularly black) students than it is for white students. For instance, among all 9th-12th graders, the residual sorting for black students is 1.4%-points compared to 0.9%-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 (Long and Conger, Forthcoming; Corcoran and Jennings 2011), 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. vii

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. viii 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 high schools appear 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.

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 the table. 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; 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.

[Insert Table 4 here]

DISCUSSION AND CONCLUSION

This study tackles an under-explored 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.

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. While 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 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, Long and Conger (Forthcoming) 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. 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.

There is also still 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. Research that aims to uncover 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.
REFERENCES


### Table 1: Summary Statistics by Gender, Graduates from Florida Public High Schools

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>All</th>
<th>Female</th>
<th>Male</th>
<th>Difference (Male-Female)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4-Year College Enrollment</strong></td>
<td>32.3%</td>
<td>35.5%</td>
<td>28.8%</td>
<td>-6.7% ***</td>
</tr>
<tr>
<td><strong>Pre-High School Achievement and Demographics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized 8th Grade Math FCAT Score</td>
<td>0.374</td>
<td>0.368</td>
<td>0.381</td>
<td>0.013 ***</td>
</tr>
<tr>
<td>Standardized 8th Grade Reading FCAT Score</td>
<td>0.354</td>
<td>0.406</td>
<td>0.298</td>
<td>-0.108 ***</td>
</tr>
<tr>
<td>White</td>
<td>59.0%</td>
<td>57.8%</td>
<td>60.2%</td>
<td>2.4% ***</td>
</tr>
<tr>
<td>Black</td>
<td>18.5%</td>
<td>19.5%</td>
<td>17.4%</td>
<td>-2.1% ***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>18.2%</td>
<td>18.4%</td>
<td>18.0%</td>
<td>-0.4% ***</td>
</tr>
<tr>
<td>Age</td>
<td>18.42</td>
<td>18.37</td>
<td>18.47</td>
<td>0.10 ***</td>
</tr>
<tr>
<td>Eligible for Free- or Reduced-Price Lunch</td>
<td>39.5%</td>
<td>40.1%</td>
<td>39.0%</td>
<td>-1.1% ***</td>
</tr>
<tr>
<td>English Language Learner (ELL)</td>
<td>13.7%</td>
<td>13.6%</td>
<td>13.7%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Special Needs</td>
<td>11.7%</td>
<td>8.2%</td>
<td>15.5%</td>
<td>7.4% ***</td>
</tr>
<tr>
<td><strong>Number of Students</strong></td>
<td>536,985</td>
<td>277,427</td>
<td>259,558</td>
<td></td>
</tr>
</tbody>
</table>

Notes:  
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) College enrollment measured as of 5 years after high school entry.  
iii) *p<.10; **p<.05; ***p<.01.
### Table 2: Gender Sorting Across Public Schools in Florida, by Race/Ethnicity

<table>
<thead>
<tr>
<th></th>
<th>Actual Distribution of Students</th>
<th>Random Allocation of Students</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>All 9-12th Graders</td>
<td>(Mean)</td>
<td>(S.D.)</td>
<td>(S.D.)</td>
</tr>
<tr>
<td>All Students</td>
<td>697</td>
<td>50.5%</td>
<td>3.4%</td>
</tr>
<tr>
<td>White Students</td>
<td>637</td>
<td>50.5%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Black Students</td>
<td>682</td>
<td>50.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>Hispanic Students</td>
<td>851</td>
<td>50.8%</td>
<td>3.2%</td>
</tr>
<tr>
<td>On-Time Graduates</td>
<td>(Mean)</td>
<td>(S.D.)</td>
<td>(S.D.)</td>
</tr>
<tr>
<td>All Students</td>
<td>441</td>
<td>48.4%</td>
<td>4.0%</td>
</tr>
<tr>
<td>White Students</td>
<td>427</td>
<td>48.9%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Black Students</td>
<td>399</td>
<td>47.2%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Hispanic Students</td>
<td>518</td>
<td>48.2%</td>
<td>3.9%</td>
</tr>
</tbody>
</table>

Notes: 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 1000 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.
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>
</tbody>
</table>

**Panel B: Decomposition Results Using HS Fixed Effects:**

|                       |                |                |                |                |
| P-Value for Test of Joint Significance of High School Fixed Effects | 0.000 | 0.000 | 0.000 | 0.000 |
| Share of Gender Gap Attributed to High School Fixed Effects | 10.9% *** | 5.2% *** | 15.8% *** | 12.2% *** |

Notes: 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>Excluding GED Recipients</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS Fixed Effect Based on Enrollment in:</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Regular HS</td>
<td>Regular HS</td>
<td>Any HS</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>
</tbody>
</table>

### Panel A: Descriptive Statistics:

- Number of Students: 536,233, 502,384, 541,637
- Male Share of Students: 48.3%, 47.7%, 48.4%
- Number of High Schools: 436, 433, 615
- Share that Attended a 4-year College:
  - Males: 28.9%, 30.7%, 28.8%
  - Females: 35.5%, 37.1%, 35.5%
  - Difference: -6.7%, -6.4%, -6.7%
  - Odds Ratio: Male Odds / Female Odds: 0.74, 0.75, 0.74

### Panel B: Decomposition Results Using HS Fixed Effects:

- P-Value for Test of Joint Significance of High School Fixed Effects: 0.000, 0.000, 0.000
- Share of Gender Gap Attributed to High School Fixed Effects: 10.9% ***, 11.5% ***, 11.8% ***

Notes: 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.
To isolate the portion of the total sorting that is due to non-random and non-mechanical influences, Long and Conger (Forthcoming) use Monte Carlo simulations that randomly allocate students to schools and they report the standard deviation in male share that arises from these simulations.

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, and 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.

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: \( \bar{Y}_f - \bar{Y}_m = \left\{ \sum_{i=1}^{N_f} \frac{F(\alpha_m X_i \delta_m + H_i \gamma_m)}{N_f} \right\} + \left\{ \sum_{i=N_f+1}^{N} \frac{F(\alpha_f X_i \delta_f + H_i \gamma_f)}{N_f} \right\} - \left\{ \sum_{i=1}^{N_f} \frac{F(\alpha_m X_i \delta_m + H_i \gamma_m)}{N_f} \right\}, \)

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 while the remainder are male.
Long and Conger (Forthcoming) and Conger (2005) use a similar method to determine the statistical significance of sorting across schools and across classrooms within-schools.

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 paper, 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 paper 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.