JANUARY 2023

Center on Children and Families at BROOKINGS

COLLEGE ENROLLMENT DISPARITIES

UNDERSTANDING THE ROLE OF ACADEMIC PREPARATION

Sarah Reber and Ember Smith

DISCLOSURES

The Brookings Institution is financed through the support of a diverse array of foundations, corporations, governments, individuals, as well as an endowment. A list of donors can be found in our annual reports, published online. The findings, interpretations, and conclusions in this report are solely those of its author(s) and are not influenced by any donation.

I. INTRODUCTION

In the U.S., a person who completes a bachelor's degree will earn about a million dollars more in their lifetime compared to someone with only a high school diploma (Carnevale et al., 2021). College graduates have higher earnings, better health, more stable marriages, and are less likely to be unemployed (Oreopoulos and Salvanes, 2011). The share of young adults with a bachelor's degree or higher doubled over the last fifty years, from 19 to 40%.[1] Despite this progress, bachelor's degree attainment still varies substantially by gender, race and ethnicity, and family socioeconomic status.

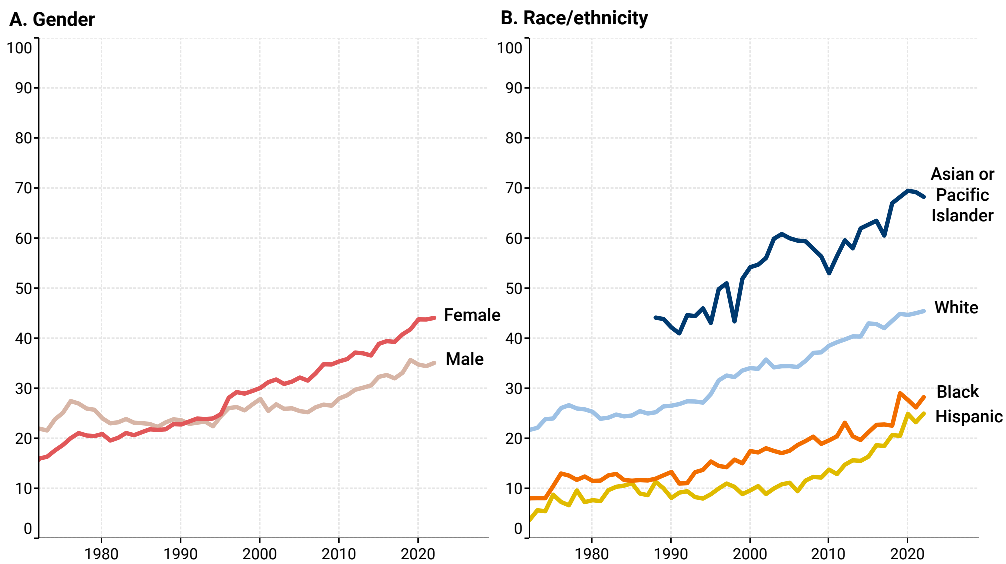
As Figure 1 shows, the gender gap in bachelor's degree attainment reversed in the last half century. In 1972, men aged 25 to 29 were six percentage points more likely to have a bachelor's degree than young women (22% compared with 16%). In 2022, young men were nine percentage points *less* likely to have a bachelor's degree than young women (35% and 44%).

Racial and ethnic disparities in attainment, by contrast, have followed a similar pattern for decades. Asian or Pacific Islander young adults[2] are the most likely to have a bachelor's degree in 2022. Disparities in bachelor's degree attainment by race and ethnicity are large: 68% of Asian or Pacific Islander adults aged 25 to 29 have a bachelor's degree, compared with 45% of white, 28% of Black, and 25% of Hispanic young adults.

FIGURE 1

Bachelor's degree attainment by gender and race/ethnicity

Percent of 25-29-year-olds with a bachelor's degree or higher, 1972 - 2022



Note: The CPS offered one combined category for Asian and Pacific Islander respondents until 2003. For comparability, we construct the same category in years after 2003. Samples for racial and ethnic groups not shown in Figure 1 were too small for estimation.
Source: Authors' calculations, Current Population Survey (CPS).

BROOKINGS

Socioeconomic disparities in college-going are more challenging to measure over time since most datasets capturing college degree attainment, including the Current Population Survey (CPS) we use in Figure 1, record a person's gender and race or ethnicity but not the characteristics of the household they grew up in. Using data from the National Longitudinal Surveys of Youth, Bailey and Dynarski (2011) show that differences in bachelor's degree attainment by socioeconomic status are large and have grown in recent years.

The root causes of these college enrollment and completion disparities are complex and differ across groups; the underlying causes of gender disparities, for instance, may be different from the causes of racial disparities, which may be different from the causes of disparities by socioeconomic status (SES). Different aspects of students' identity may also interact with each other; for example, gender disparities may differ by SES or race. Despite this limitation, we think that identifying proximate causes of socioeconomic, gender, and racial enrollment gaps is a valuable starting point for understanding root causes and designing interventions to address them.

In this report, we evaluate the role of differences in academic experiences and skills—which we refer to as "academic preparation" and measure with test scores, high school grades, and course-taking—as a proximate cause of college enrollment disparities. We use the restricted-use High School Longitudinal Survey (HSLS 2009) dataset, a nationally representative sample of students who were in ninth grade during the 2009-10 school year, to estimate the gender, racial, and socioeconomic gaps in college enrollment overall and among students with similar academic preparation. We also assess the extent to which racial differences in socioeconomic status, alone or in combination with the academic preparation measures, account for racial differences in college enrollment outcomes.

We find that academic preparation explains a substantial portion of socioeconomic, gender, and racial gaps in college enrollment. SES gaps in enrollment, especially in four-year colleges, are enormous; 89% of students from families in the top SES quintile enroll in college compared with 51% of those in the bottom—a 38-point gap. Among students with similar academic preparation, the gap is much smaller though still notable. Students from the top quintile are 11 points more likely to enroll in college than students from the bottom quintile who have similar academic preparation.

As previous work finds, girls perform better in school than boys and are more likely to enroll in college (Goldin, 2006; Reeves, 2022). We find that 73% of girls and 64% of boys enroll in college but that girls and boys with the same academic preparation enroll in college at about the same rate. In other words, the gender difference in high school academic preparation fully accounts for the gender gap in college enrollment.

Enrollment rates also differ significantly by race or ethnicity: 83% of Asian, 72% of white, 63% of Hispanic, and 62% of Black students enroll in college within a year and a half of expected high school graduation. Differences in academic GPA, course-taking, and test scores follow a similar pattern. Among students with the same high school academic preparation, Black, Hispanic, and Asian students all enroll in college at about the same rate—five points higher than the rate for white students. Among students with similar socioeconomic status, Asian students enroll in any college at the highest rate, and enrollment rates are similar for Black, Hispanic, and white students.

While college admissions and cost are significant factors in discussions about unequal access to higher education, our findings indicate that policymakers and researchers should also focus on disparities in academic preparation during elementary and secondary education. Differences in academic preparation are not only influenced by a student's actions but also by the opportunities they have to learn both in and out of school, and group differences in academic preparation could be the result of discrimination affecting a student's opportunities to learn. Closing gaps in academic preparation is crucial for addressing gender and racial gaps in college enrollment, and addressing both academic and non-academic factors like cost and lack of information will be necessary to address socioeconomic disparities among students with similar levels of academic preparation.

II. DATA

We use the High School Longitudinal Survey (HSLS), the most recent longitudinal survey conducted by the National Center for Education Statistics (NCES). The survey includes over 23,000 students who were in ninth grade in 2009.[3] Students were surveyed several times throughout their time in high school and early adulthood and took a standardized math exam, administered by NCES, in their expected ninth and eleventh grade years. The survey records were linked to data from high school transcripts and college enrollment records, as well as surveys completed by the student's parent or guardian, school administrators, counselors, and teachers. The longitudinal design of the HSLS uniquely allows us to connect college outcomes, like enrollment, to information about students' high school experience and family background. We restrict the sample to just over 15,000 students who have non-missing data for the outcome and control variables described below.

College enrollment measures

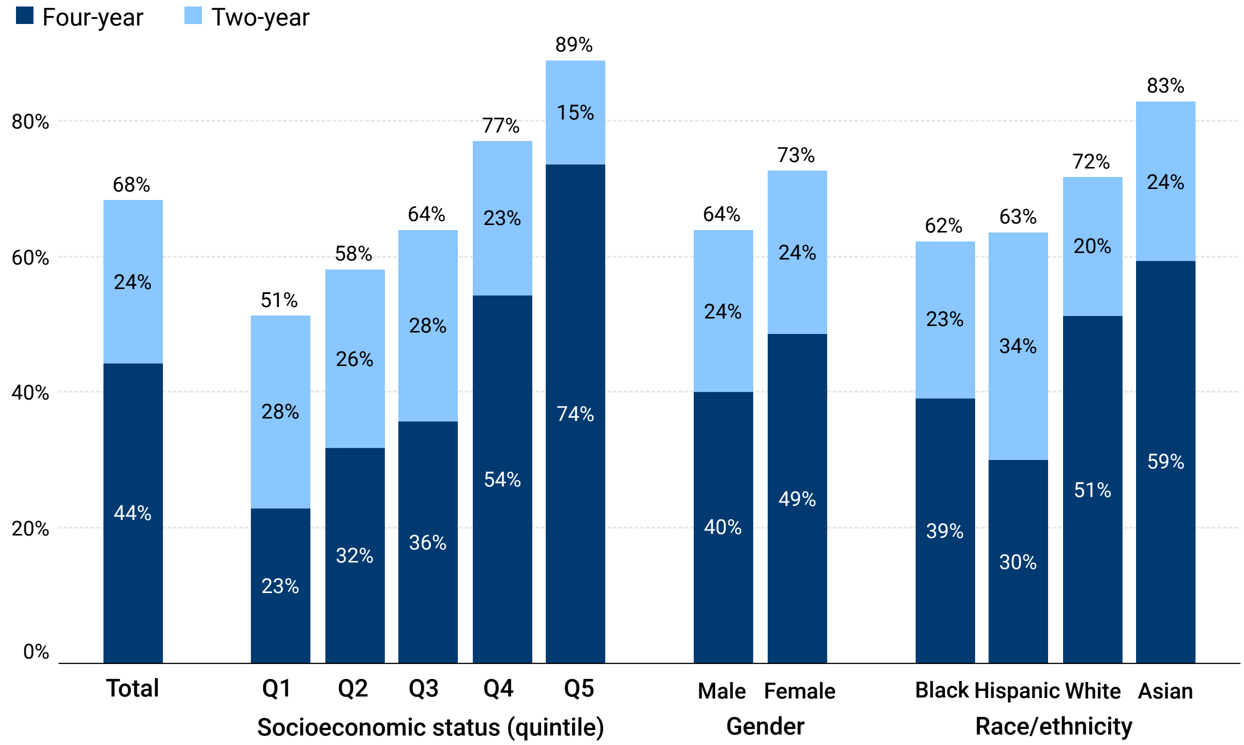
The key outcome of interest is enrollment in postsecondary education and whether enrollment was in a two- or four-year institution. We assign students to one of three mutually exclusive categories based on the type of institutions in which they enrolled at any point during the first 18 months after expected high school graduation[4]:

- *Four-year* includes institutions that grant bachelor's degrees.
- *Two-year* includes institutions that grant associate or other two-year degrees; institutions that only grant certificates or one-year degrees are excluded.
- *No college* includes people who never enrolled in a two-year or four-year college during the 18 months following expected high school graduation; note this classification includes people who enroll in institutions that only grant certificates.

FIGURE 2

Postsecondary enrollment rate

Percent of 2009 9th graders enrolled within 18 months of expected HS graduation



Source: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09). See text and Table A1 for details.

BROOKINGS

We assign students who enrolled in more than one type of institution to the group corresponding to the highest-level institution they attend. That is, if a student enrolled in both a two-year and a four-year college within 18 months of their expected high school graduation date, we assign them to the four-year category.[5] Similarly, if a student dropped out of college after a semester, we assign them to the highest institution level they attended at any point within 18 months of expected high school graduation. Figure 2 presents summary statistics for these outcomes by gender, race, and socioeconomic status. In our full sample (first bar), 68% of students enrolled in college: 44% enrolled in a four-year college (the dark blue bar) and 24% in a two-year college (the light blue bar). See Appendix Table A1 for estimates with confidence intervals and counts rounded to protect respondent privacy. We discuss the summary statistics for each of the three demographic categories below.[6]

Socioeconomic status, gender, and race or ethnicity

NCES coded sex (which we refer to as gender throughout) into two categories—male and female—based on the ninth-grade student questionnaire. If a respondent's sex is missing, it is supplemented with parent or school surveys.[7] We use quintiles of socioeconomic status (SES) constructed by the NCES based on parent or guardians' education, occupation, and income.[8] Finally, we rely on mutually exclusive categories constructed by NCES to classify students by race and ethnicity (which we refer to as "race" going forward for expositional ease): American Indian or Alaska Native (AIAN); Asian; Black; Hispanic; more than one race; Native Hawaiian or Pacific Islander (NHPI); and white.[9]

We include all students in the analytic sample and report results for each race/ethnicity in Appendix Tables A1 and A4, but we do not discuss the findings for AIAN or NHPI students (because the samples are too small to yield reliable estimates) or multi-racial students (because these findings are difficult to interpret).

Academic preparation

We use several variables drawn from students' high school transcripts and the score on a standardized math test to measure academic preparation. We chose these measures because they are highly predictive of college enrollment outcomes and measured consistently across settings, although there may be some differences in grading standards across schools.

Academic GPA is the unweighted average of grades in English, mathematics, science, social studies, foreign language, and fine arts on a standard four-point scale. We also consider *Math GPA* and *English Language Arts (ELA) GPA* separately. All GPA variables were constructed from the transcripts by NCES.[10] For all three GPA measures, we use decile indicators based on the full sample of students with reported GPAs. The rigor of the coursework a student took in high school is also important for college enrollment, but variation in curriculum and course naming conventions across high schools makes this difficult to fully characterize. Still, we create several indicators of curricular rigor that are reasonably comparable across schools: indicators for whether the student took advanced math and highly advanced math and the total number of Advanced Placement (AP) or International Baccalaureate (IB) courses a student took.[11]

We also use the score on a math test that was administered by NCES when most students were in eleventh grade. The math assessment covered algebraic reasoning using an adaptive two-stage format to reduce floor and ceiling effects. We standardize the math score to have a mean of zero and a standard deviation of one in the full sample and use decile indicators in the regression analyses described below.[12] NCES did not administer an ELA test as part of HSLS. The math test score has the advantage of being standard across settings, but it measures a narrower set of academic skills than the school-based measures since it only captures math performance through algebra. It is also less reflective of non-cognitive skills important for post-secondary outcomes that are more likely be captured by students' grades than test scores.

Groups with higher college enrollment rates also tend to have higher GPAs, test scores, and have taken a more rigorous course of study in high school.[13] We discuss the summary statistics for these variables in detail separately by socioeconomic status, gender, and race in the results section below.

III. METHODS

Academic preparation is, not surprisingly, strongly predictive of college enrollment.[14] Average academic preparation also varies considerably by socioeconomic status, gender, and race. This raises a simple question: Do the differences in academic preparation explain the differences in college enrollment outcomes by socioeconomic status, gender, and race?

To assess the extent to which academic preparation disparities are responsible for college enrollment disparities, we use a simple linear regression framework. This approach allows us to consider whether students with similar levels of academic preparation but different socioeconomic status, gender, or race have similar college enrollment outcomes. We estimate linear regressions of the following form—illustrated here for our analysis by socioeconomic status, but we take the same approach for analyses by gender and race:

$$y_i = a + \beta_1 \times SESQ1_i + \beta_2 \times SESQ2_i + \beta_3 \times SESQ4_i + \beta_4 \times SESQ5_i + X_i\Phi + \varepsilon_i$$

where y_i is a binary college enrollment outcome (any, four-year, or two-year enrollment), and the SESQ variables are a series of indicator variables for each SES quintile with the third quintile as the reference group, and X_i is a vector of control variables.[15] We estimate several versions of this regression with different combinations of control variables for each group (socioeconomic status, gender, and race) and outcome.[16] The specifications are as follows:

1. No controls
2. Academic GPA only
3. Math GPA only
4. ELA GPA only
5. Math test score only
6. Course-taking (advanced and highly advanced math, number of AP/IB credits)
7. All academic preparation measures (GPA, Math GPA, ELA GPA, math test score, and course-taking)
8. SES only
9. Academic preparation (all measures in specification (7)) & SES

We only estimate specifications (8) and (9) for the analysis by race. We present the main results graphically, showing the most relevant specifications for each category, and present all specifications in the appendix.

When control variables are included in the model, we can only identify the differences in outcomes between groups, conditional on the controls; the level is not identified. For the graphs, we rescale the coefficients reported in Table A2 by adding the average of the outcome for the reference group (in this example, the third quintile) to the coefficient. This allows for easier comparison of the differences across groups to the average level of the outcome and means that the level for the reference group is constant across the specifications presented in the figure.

We compare the β coefficients for specification (1), the unconditional gaps, to the β coefficients controlling for academic preparation as in, for example, specification (7). We say that group differences in the relevant outcome are "explained by" differences in academic preparation to the extent that including those controls in the regression reduces the β coefficients.[17] We do not necessarily interpret the coefficients on the academic preparation variables (Φ) causally, nor do we interpret conditional gaps (the β coefficients) as evidence of discrimination (or lack of discrimination).[18] Rather, we interpret the regressions descriptively as revealing conditional gaps, allowing us to compare outcomes for students with similar GPA, test scores, etc. The relationship between the academic preparation variables and college enrollment is almost certainly at least partially causal, but there may be unobserved student characteristics—other aspects of academic preparation, family background characteristics, or non-cognitive skills, for example—that are correlated with these measures of academic preparation and influence college enrollment. In that case, we will attribute the effect of those unobserved characteristics to the academic preparation measures we include.

We also emphasize that differences in academic preparation depend not only on a student's actions but also their opportunities to learn, both in and out of school. For example, a student might not have taken advanced math because it is not offered at their school, because they did not have strong math instruction in their early schooling, or because they did not spend enough time studying, which itself could be influenced by factors beyond the student's control. Group differences in academic preparation could result from discrimination, ongoing or historical, affecting students' opportunities to learn through a wide range of channels. And academic preparation could be influenced by students' expectations about their access to college; if they do not think they will be able to attend college for financial reasons, for example, they might not spend as much time studying in high school. This analysis cannot speak to why or how differences in academic preparation arise. Despite these limitations, because academic preparation contributes to whether a young person enrolls in college, we think assessing the magnitude of enrollment disparities conditional on academic preparation is valuable. For example, if SES gaps in enrollment are largely explained by differences in academic preparation, policymakers should prioritize closing academic preparation gaps. Otherwise, targeting interventions to other factors, such as lack of information or difficulty financing college might be promising.

IV. RESULTS:
DO ACADEMIC PREPARATION GAPS EXPLAIN COLLEGE ENROLLMENT GAPS?

In this section, we consider each of three categories—socioeconomic status, gender, and race—in turn. For each group, we begin by presenting summary statistics for the college enrollment and high school performance measures before presenting the analysis of college enrollment disparities.

Socioeconomic Status

Table 1 presents the summary statistics for the college enrollment outcomes (panel a) and the measures of academic preparation (panel b), by quintiles of socioeconomic status. The final column shows the averages for the full sample. We also show the distribution of the sample across the other groups, in this case gender and race (panel c).

Overall, about 44% of students enroll in a four-year college, 24% in a two-year college, and 32% do not enroll in any college in the first 18 months after expected high school graduation. The gradient by socioeconomic status in college enrollment overall and especially for four-year enrollment is steep: About half of students in the bottom SES quintile enroll in college, compared to 89% of students in the top quintile. Almost three-quarters of students in the top quintile enroll in a four-year college, compared to less than a quarter in the bottom quintile. Two-year enrollment rates are similar in the bottom three quintiles, and lower in the fourth and fifth quintiles—which is more than offset by their higher four-year enrollment rates. As shown in Table 1 and the first set of bars in Figure 3, college-going is strongly associated with socioeconomic status.

Panel b of Table 1 shows that differences in all the measures of academic preparation mirror those for college enrollment outcomes. Students from higher SES households have higher GPAs and math test scores and take more rigorous coursework. Math test scores are more strongly related to socioeconomic status than math GPA, which may relate to differences in advanced course-taking. For example, 43% of students in the top fifth of the SES distribution take highly advanced math, compared to 22 or 23% in the bottom three quintiles. Taking more advanced math may help students do better on the standardized test more than it helps their GPA, depending on grading standards. It is also possible the causality goes the other way; students with high math scores may be placed in advanced math.

Figure 3 presents the key results for socioeconomic status for any college enrollment (panel a) and four-year college enrollment (panel b). The first set of bars—1. No controls—presents the unconditional means (as in panel a of Table 1 and column 1 of Table A2). The second set

TABLE 1

College enrollment and academic performance by socioeconomic status

	First quintile (lowest)	Second quintile	Third quintile	Fourth quintile	Fifth quintile (highest)	Total
A. College enrollment outcomes						
Enrollment (%)						
Any college	51.2	58.0	63.9	76.9	88.9	68.3
Four-year	22.9	31.7	35.7	54.3	73.5	44.3
Two-year	28.2	26.3	28.1	22.6	15.3	24.0
B. Academic performance						
Academic GPA						
Mean (4-pt scale)	2.22	2.39	2.49	2.71	3.06	2.59
Bottom decile (%)	11.6	9.5	5.4	3.6	2.0	6.3
Top decile (%)	4.0	6.7	7.9	14.1	23.6	11.5
ELA GPA						
Mean (4-pt scale)	2.25	2.40	2.52	2.73	3.06	2.60
Bottom decile (%)	13.6	10.8	7.0	4.4	2.3	7.5
Top decile (%)	3.9	7.5	8.6	13.4	22.5	11.4
Math GPA						
Mean (4-pt scale)	1.99	2.15	2.24	2.46	2.82	2.34
Bottom decile (%)	13.6	11.2	6.9	4.8	2.5	7.7
Top decile (%)	4.5	6.9	8.3	13.4	21.8	11.2
Math exam						
Standardized z-score	-0.35	-0.16	-0.05	0.28	0.72	0.10
Bottom decile (%)	15.6	10.8	8.7	5.7	2.5	8.5
Top decile (%)	3.9	5.8	6.9	14.0	27.2	11.9
Math course-taking (%)						
Advanced math or higher	46.6	51.1	52.7	65.1	81.2	59.8
Highly advanced math	21.6	23.5	23.5	30.9	42.9	28.8
AP/IB course-taking (%)						
1 or more	20.6	26.4	28.0	39.0	59.4	35.2
3 or more	8.9	10.4	13.0	19.5	34.6	17.6
6 or more	1.9	3.3	3.7	7.4	13.9	6.2
8 or more	0.9	0.9	1.2	3.0	6.5	2.6
C. Demographic						
Gender (%)						
Male	52.5	49.0	47.7	51.1	50.0	50.1
Female	47.5	51.0	52.3	48.9	50.0	49.9
Race (%)						
Asian	2.7	2.8	2.4	3.7	6.4	3.6
Black	18.1	16.3	14.0	10.6	5.3	12.7
Hispanic	41.7	26.0	20.2	12.9	8.2	21.5
White	29.1	46.6	53.1	64.0	72.1	53.5
N	**2,170**	**2,510**	**2,790**	**3,330**	**4,300**	**15,090**

NOTE: Counts are rounded to the nearest ten to protect respondent privacy. See text for details on the sample and weights.
SOURCE: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09).

of bars—2. GPA only—shows how college enrollment differs across SES group, controlling for deciles of academic GPA only, and the third set of bars—7. Academic preparation—show how enrollment differs controlling for all the academic preparation measures.[19]

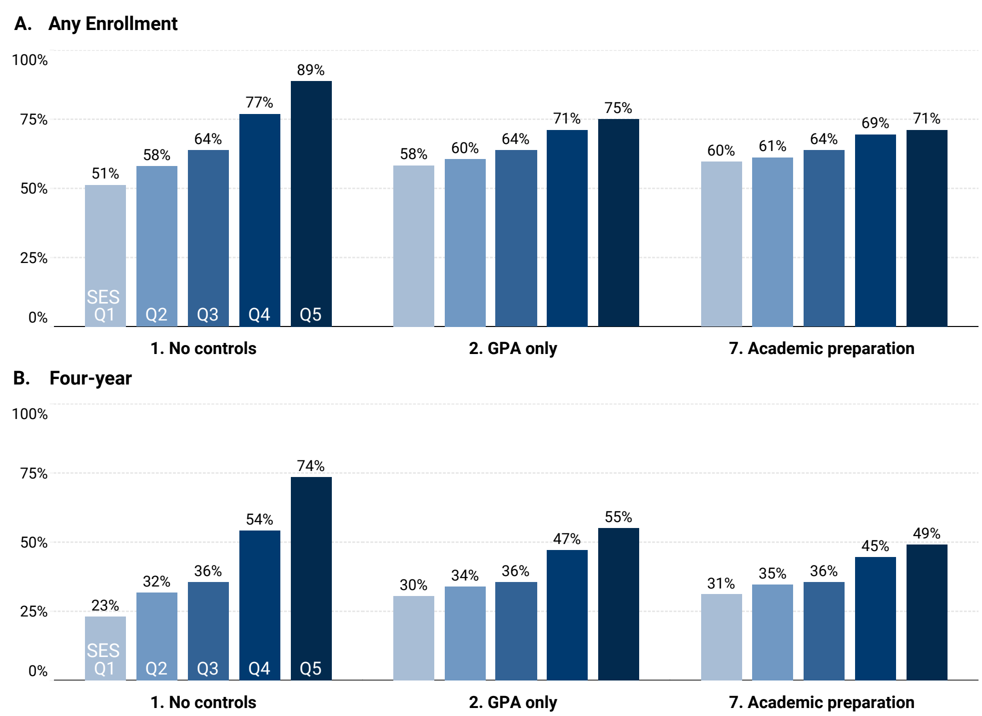
As discussed above, the SES gradient in any college enrollment is steep, with the gap increasing especially at the fourth and fifth quintiles. The second set of bars shows that controlling for academic GPA reduces the gradient considerably but does not eliminate them. The gap between the top and bottom quintiles falls from 38 to 16 percentage points, implying that academic GPA accounts for about 56% of the gap in any college enrollment between those two groups. Academic GPA accounts for a similar share (57%) of the gap between the first and third quintiles. The gradient is reduced further in the third set of bars, where we control for all the academic preparation variables. Academic GPA, math test scores, and course-taking account for about 70% of the fifth to first quintile gap and of the third to first quintile gap. These results suggest that enrollment gaps by socioeconomic status are substantially smaller, though still notable, among students with similar academic preparation.

Panel b shows the same results with four-year enrollment as the outcome. The patterns in four-year enrollment are similar to any college enrollment but more pronounced, with the top two quintiles especially likely to enroll in a four-year college. Differences in academic preparation explain about 65% of the 51 percentage-point gap between the first and fifth quintile. Even among those who have similar academic preparation, students in the top SES quintile are 18

FIGURE 3

Postsecondary enrollment rate, by socioeconomic status (SES) quintile

Percent of 2009 9th graders enrolled within 18 months of expected HS graduation



Source: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09). See text and Table A2 for details.

BROOKINGS

percentage points more likely to enroll in a four-year college than students in the bottom SES quintile.

For all three of the college enrollment outcomes, the results from specifications that include just math GPA deciles, just ELA GPA deciles, just math test deciles, or just the course-taking variables follow a similar pattern to the estimates controlling for academic GPA deciles alone but with a slightly steeper unexplained SES gradient. That is, any of the GPA variables alone can account for a significant share of the SES gradient, but academic GPA accounts for the most on its own, and all the variables together account for slightly more than GPA alone (see Table A2). Measures of academic preparation are highly correlated with each other, so having information on any of them goes a long way (but not all the way) in explaining enrollment gaps.

Overall, our results show that students from socioeconomically advantaged families are more likely to enroll in college and especially likely to enroll in a four-year institution relative to their peers. Much of this pattern is explained by differences in academic preparation, but even among students with similar academic preparation, gaps in enrollment—especially at four-year colleges—are large.[20]

Gender

Table 2 shows the summary statistics by gender; the final column shows averages for the full sample and is the same as the final column in Table 1. About 27% of women did not enroll in any college, compared to 36% of men. This gap is driven by the difference in four-year enrollment: About 24% of both women and men enroll in a two-year college, but 49% of women, compared to 40% of men enroll in a four-year institution, accounting for the nine-percentage point gap in overall enrollment.

Whereas each of the measures of academic preparation tells essentially the same story about differences by SES, differences by gender vary more across measures. Boys do relatively better on tests than measures of performance in school like GPA. The average academic GPA for female students is 0.32 grade points higher than for male students and there is no gender gap in the average math test score, although males are more likely to be in both the top and bottom deciles; this “males in the tails” phenomenon has been well-documented (Autor et al., 2020). Boys achieve similar scores on the math test despite having somewhat lower rates of advanced math course-taking than girls.

About 24% of both women and men enroll in a two-year college, but 49% of women, compared to 40% of men enroll in a four-year institution

NCES did not administer a test in subjects other than math, so we cannot determine if the GPA-test score divergence is because boys are relatively better at math compared to other subjects (a “subject effect”) or because they are relatively better at taking tests than getting good grades (a “test effect”). The evidence points to some role for both. The gender gap favoring girls in math GPA is smaller than for ELA GPA (0.23 points compared to 0.38 points); that is, girls do better in both subjects, but less so in math. This could be partly attributable to girls’ higher-level course-taking; for example, girls are about six percentage points more likely to take advanced math, which could reduce their GPA advantage relative to other subjects. Other research finds that girls typically outperform boys by a larger margin on English exams relative to math exams (Reardon et al., 2019). On math tests, there is typically no gender gap or even a gap favoring males. Overall, both a “test effect” and a “math effect” appear to be at play in explaining the fact that the math test score has less explanatory power for the gender gap in postsecondary enrollment than other measures of academic preparation.

Figure 4 shows the results of analyzing gender gaps; we include more specifications than in Figure

TABLE 2

College enrollment and academic performance by gender

	Male	Female	Total
A. College enrollment outcomes			
Enrollment (%)			
Any college	64.0	72.6	68.3
Four-year	40.0	48.6	44.3
Two-year	23.9	24.0	24.0
B. Academic performance			
Academic GPA			
Mean (4-pt scale)	2.43	2.75	2.59
Bottom decile (%)	8.7	3.9	6.3
Top decile (%)	7.9	15.2	11.5
ELA GPA			
Mean (4-pt scale)	2.41	2.80	2.60
Bottom decile (%)	10.4	4.5	7.5
Top decile (%)	7.2	15.6	11.4
Math GPA			
Mean (4-pt scale)	2.23	2.46	2.34
Bottom decile (%)	9.8	5.5	7.7
Top decile (%)	8.9	13.5	11.2
Math exam			
Standardized z-score	0.12	0.09	0.10
Bottom decile (%)	9.7	7.3	8.5
Top decile (%)	13.3	10.4	11.9
Math course-taking (%)			
Advanced math or higher	56.9	62.6	59.8
Highly advanced math	27.8	29.8	28.8
AP/IB course-taking (%)			
1 or more	30.3	40.1	35.2
3 or more	15.4	19.9	17.6
6 or more	5.7	6.7	6.2
8 or more	2.6	2.6	2.6
C. Demographic			
Socioeconomic status (%)			
First quintile (lowest)	20.5	18.5	19.5
Second quintile	18.6	19.4	19.0
Third quintile	18.9	20.8	19.8
Fourth quintile	20.8	20.0	20.4
Fifth quintile (highest)	21.3	21.3	21.3
Race (%)			
Asian	3.7	3.6	3.6
Black	11.1	14.3	12.7
Hispanic	21.4	21.5	21.5
White	55.5	51.4	53.5
N	**7,320**	**7,770**	**15,090**

NOTE: Counts are rounded to the nearest ten to protect respondent privacy. See text for details on the sample and weights.

SOURCE: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09).

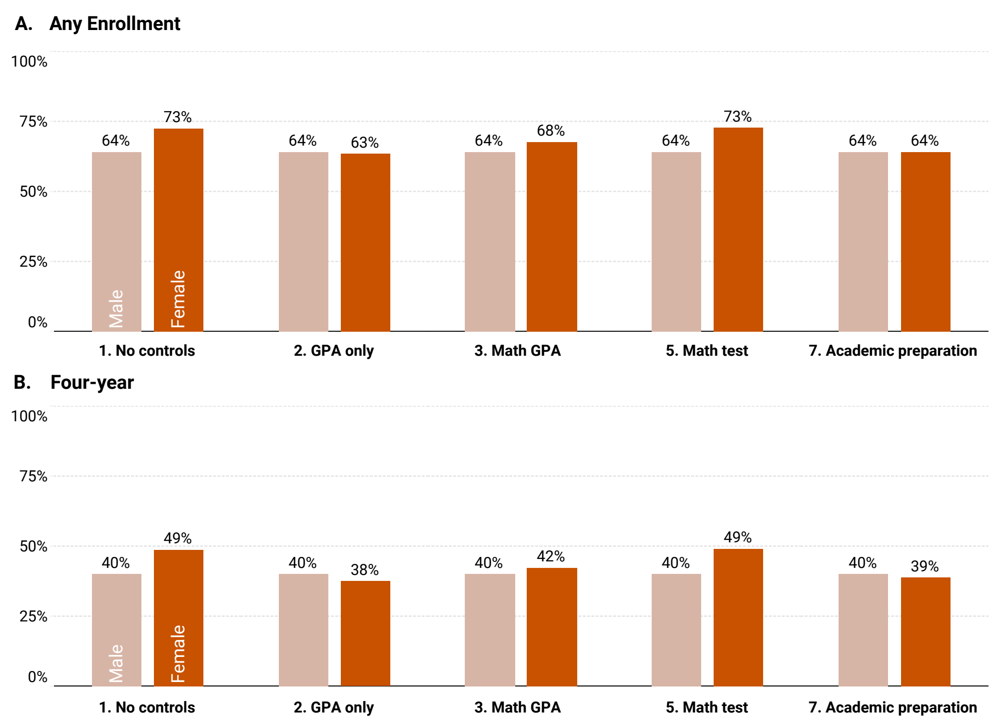
3 because the choice of measure matters more for gender than SES. Gender differences in high school GPA entirely account for the nine-point gender difference in any postsecondary enrollment (panel a). In other words, women and men with similar GPAs enroll in college at similar rates. When accounting for all academic preparation, gender gaps look like those when accounting for only academic GPA. Controlling for any measure of academic preparation has a similar impact on the conditional SES gaps, but the same is not true for conditional gender gaps. The full results reported in Table A3 show that the gender gaps are almost fully explained in specifications that account only for academic GPA, only for ELA GPA, and all academic preparation. Controlling for the math score does not reduce the conditional gender gap at all, which makes sense because there is no gender gap in math scores in this sample.[21] Controlling for math GPA explains about 60% of the gender gap, compared to slightly more than 100% for ELA GPA (meaning boys enroll in college more than girls with similar ELA GPAs) and 0% for the math test score.

The pattern of results is similar for four-year enrollment and in some cases the gender gaps reverse when controls for academic preparation are included. The second set of bars shows that among students with similar GPAs, four-year college attendance is about 2.4 percentage points lower for girls than boys. In other words, boys enroll at a higher rate than girls with the same GPA. Controlling only for the math score, on the other hand, increases the conditional gap, though the magnitude is small (see Table A3). Accounting for all academic preparation measures, girls are about 1.1 percentage points less likely to attend a four-year college

FIGURE 4

Postsecondary enrollment rate, by gender

Percent of 2009 9th graders enrolled within 18 months of expected HS graduation



Source: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09). See text and Table A3 for details.

BROOKINGS

than boys, though this difference is not statistically significant. Among students with similar academic preparation, boys are somewhat more likely to enroll in a four-year college and girls in a two-year college, but these differences are small and sometimes not significant. Taken together, the results by gender suggest that most or all gender gaps in college enrollment are explained by differences in academic preparation. However, in contrast to the results by SES, it is important which measure of academic preparation is considered; GPA explains essentially all, and math test score explains none, of the gaps.

Race

In this section, we examine racial disparities in college enrollment and evaluate the extent to which these disparities are present among students with the same academic preparation, as we did for socioeconomic status and gender above. We also describe differences in socioeconomic status by race and evaluate racial gaps in enrollment among students with enrollment among students with similar socioeconomic status.

College enrollment outcomes vary substantially across racial groups. Table 3 shows summary statistics by race for the four largest racial categories; unfortunately, we do not have large enough samples to analyze outcomes for American Indian or Alaska Natives (AIAN) or Native Hawaiian or Pacific Islanders (NHPI) separately.[22] Asian students are the most likely to enroll in any college (83%), followed by white students (72%), and Hispanic and Black students (63% and 62%, respectively). The pattern is broadly similar for four-year college enrollment. Hispanic students are much more likely to enroll in a two-year college compared to Black students (34% versus 23%), even though both groups have similar rates of overall college enrollment. The measures of academic preparation follow a similar pattern to the college enrollment outcomes by race: The groups more likely to enroll in college also have stronger academic preparation in high school across all the measures.

Next, we analyze college enrollment by race among students with similar academic preparation. The first set of bars in Figure 5 shows the unconditional gaps in any college enrollment.[23] The highest-enrollment group (Asian students) is more than 20 percentage points more likely to enroll in college than the lowest-enrollment group (Black students). Asian students are 11 percentage points more likely to enroll than white students, and 19 percentage points more likely than Hispanic students. Hispanic students are about one percentage point more likely to enroll in any college than Black students, though this difference is not statistically significant.

The third set of bars shows results controlling for all measures of academic preparation. Among students with similar academic preparation, Black students are most likely to enroll in college (seven and eight percentage points more than white or Asian students), followed by Hispanic students (four and five percentage points more than white or Asian students). White and Asian students with similar academic preparation enroll in any college as similar rates. The enrollment pattern after accounting for academic GPA (shown in second set of bars) falls in between the unconditional disparities and the specification including all of the academic preparation controls.[24]

Unconditional four-year enrollment rates (panel b) follow a similar pattern as enrollment in any college, though some of the gaps are even larger, especially compared to the average. Conditional on all academic preparation measures, Asian students are the least likely to enroll in four-year college—about 18 percentage points less likely than Black students, who have the highest rate of four-year college enrollment. Our results suggest that, among students with similar academic preparation: Black students are the most likely to enroll in college, especially four-year college; Hispanic students are almost as likely as Black students to enroll in any college but more likely to enroll in a two-year institution; and Asian and white students enroll in any college at similar rates, but Asian students are substantially less likely to enroll in a four-year college.

While racial enrollment gaps are smaller or reversed among students with similar academic preparation, a student's opportunities to learn also depend on their

TABLE 3

College enrollment and academic performance by race/ethnicity

	Asian	White	Black	Hispanic	Total
A. College enrollment outcomes					
Enrollment (%)					
Any college	82.9	71.7	62.2	63.5	68.3
Four-year	59.3	51.2	39.2	30.0	44.3
Two-year	23.6	20.5	23.0	33.5	24.0
B. Academic performance					
Academic GPA					
Mean (4-pt scale)	3.01	2.78	2.21	2.31	2.59
Bottom decile (%)	1.0	4.2	12.4	9.3	6.3
Top decile (%)	20.2	16.5	3.3	4.0	11.5
ELA GPA					
Mean (4-pt scale)	3.03	2.78	2.23	2.34	2.60
Bottom decile (%)	2.7	5.8	12.7	9.8	7.5
Top decile (%)	19.3	16.1	3.3	4.2	11.4
Math GPA					
Mean (4-pt scale)	2.81	2.55	1.99	2.02	2.34
Bottom decile (%)	1.1	5.1	13.1	11.5	7.7
Top decile (%)	20.6	15.2	3.9	5.2	11.2
Math exam					
Standardized z-score	0.88	0.27	-0.38	-0.12	0.10
Bottom decile (%)	2.2	6.5	15.6	10.1	8.5
Top decile (%)	37.3	14.5	2.5	7.4	11.9
Math course-taking (%)					
Advanced math or higher	83.6	63.5	59.8	50.1	59.8
Highly advanced math	53.4	30.5	22.7	24.9	28.8
AP/IB course-taking (%)					
1 or more	68.7	37.5	23.5	31.8	35.2
3 or more	47.1	18.5	10.1	15.9	17.6
6 or more	21.4	6.1	3.0	5.7	6.2
8 or more	11.4	2.4	0.8	2.7	2.6
C. Socioeconomic status (%)					
Gender (%)					
Male	51.0	52.0	43.8	49.9	50.1
Female	49.0	48.0	56.2	50.1	49.9
Race (%)					
First quintile (lowest)	14.4	10.6	27.8	37.8	19.5
Second quintile	14.6	16.6	24.4	23.0	19.0
Third quintile	12.8	19.7	21.9	18.7	19.8
Fourth quintile	20.9	24.5	17.1	12.3	20.4
Fifth quintile (highest)	37.3	28.7	8.8	8.1	21.3
N	**1,300**	**8,590**	**1,460**	**2,270**	**15,090**

NOTE: Counts are rounded to the nearest ten to protect respondent privacy. See text for details on the sample and weights.

SOURCE: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09).

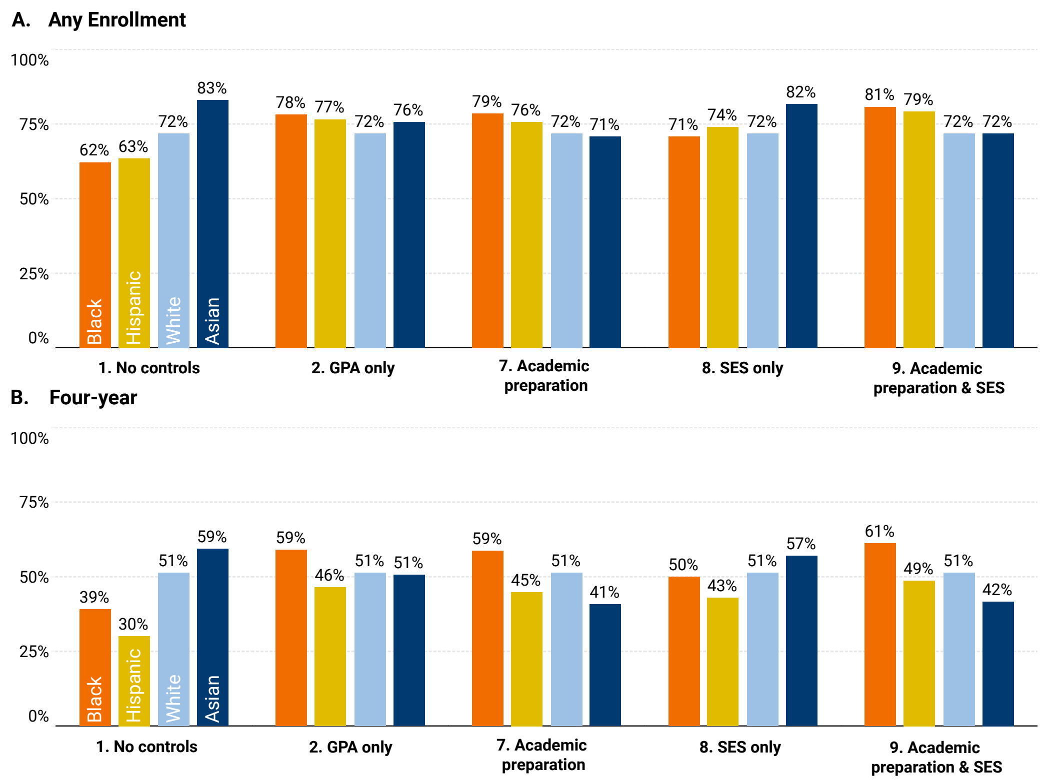
family's resources. And, not surprisingly given patterns of immigration and the history of discrimination against Black people and other people of color in the United States, students' SES varies considerably by race (Table 3, panel c). Asian students are most likely to be in the top quintile of SES (37%), and Hispanic students are most likely to be in the bottom quintile (38%). To estimate the extent to which the differences in college enrollment outcomes by race can be explained by differences in SES, we include two additional specifications for the analysis of racial disparities, in addition to the results accounting for academic preparation: controls for SES only and controls for SES and all academic preparation variables.

The fourth and fifth set of bars in Figure 5 show how controlling for SES alone or together with the academic preparation measures influences the conditional gaps. Controlling for SES alone accounts for most of the gaps in any college enrollment among Black, Hispanic, and white students, and Asian students are about 10 percentage points more likely to enroll in college conditional on SES. However, controlling for SES in addition to the academic preparation measures does not change the picture too much, suggesting that the SES effects likely operate largely through differences in academic preparation.[25] The results for four-year enrollment suggest that SES differences partially account for Hispanic students' lower four-year enrollment (recall that almost 40% of Hispanic students are in the bottom SES quintile), but the same is not true for Asian students.

FIGURE 5

Postsecondary enrollment rate, by race/ethnicity

Percent of 2009 9th graders enrolled within 18 months of expected HS graduation



Source: Authors' calculations based on the High School Longitudinal Survey of 2009 (HSLS:09). See text and Table A4 for details.

BROOKINGS

V. DISCUSSION

In this paper, we examine disparities in college enrollment outcomes by socioeconomic status, gender, and race, with a focus on the role of academic preparation in explaining these gaps. We measure academic preparation using GPA, course-taking, and math test scores. For all three categories, college enrollment gaps are large, and differences in academic preparation explain a substantial portion of the gaps, though the magnitudes differ depending on the group and outcome.

Disparities by SES are particularly large, especially for four-year college enrollment: Students in the top SES quintile are three times as likely as students in the bottom quintile to enroll in a four-year college. Over half of the gap between the top- and bottom-quintile college enrollment rates is explained by differences in academic preparation, but there are still large differences in enrollment rates among students with similar academic preparation but different SES.

Across SES, gender, and race, academic GPA alone has the most power to explain enrollment differences, though the academic preparation measures are highly correlated with each other. For SES and race, any of the academic preparation measures—like math GPA and test scores—by themselves explain a sizable share of the enrollment gap. By contrast, GPA and test score are not good substitutes when studying gender gaps; GPA alone fully explains the gender gap while gender differences in math scores alone have no explanatory power. This is consistent with previous research, which also finds that gender differences in GPA and test scores can diverge (e.g., Duckworth and Seligman, 2006; Goldin et al., 2006). Future research should focus on examining the role of behavior measures and non-cognitive skills in explaining gender gaps in educational attainment, as these factors may influence GPA more than test scores (Becker et al., 2010; Owens, 2016; Reeves, 2022). This also has implications for studies using less complete data. For example, researchers often have standardized test scores but not GPA or course-taking measures. Our findings suggest that controlling for test scores will produce similar results (though magnitudes may differ) as GPA and course-taking when examining outcomes by race or SES, but not by gender.

College enrollment rate differences by race are complex. Asian students have the highest postsecondary enrollment rates, followed by white, Hispanic, and Black students. However, among students with similar academic preparation, the order changes: Black students are the most likely to enroll in college, followed by Hispanic, white, and Asian students. Enrollment rates also differ by institution level; controlling for academic preparation, Hispanic and Asian students who attend college are more likely to attend a two-year college than white or Black students. Differences in SES are important in explaining differences in college enrollment rates by race, largely because of differences in academic preparation related to SES.

Our findings are in line with previous research on disparities by SES, gender, and race (Ellwood and Kane, 2000; Phillips, 2011; Goldin et al., 2006). Although analytic approaches and definitions of academic preparation differ, our findings, taken together with previous work, suggest consistency in both the size of enrollment gaps by SES, gender, and race in recent decades and the role academic preparation plays in explaining them. In future work, we will compare our findings for this recent cohort to earlier cohorts using comparable measures, which will help us better understand how gaps—and the contribution of differences in academic preparation to them—have changed over time.

We note several limitations of our analysis and caveats about the interpretation of our findings. First, the measures of academic preparation and SES we use may not capture all relevant factors related to college enrollment, and if unobserved factors differ across groups, our analysis may under- or overstate the contribution of academic preparation and SES to college enrollment disparities. Although the academic preparation measures we consider almost certainly have some causal relationship with college enrollment, the relationship we identify may be partly due to factors correlated with academic preparation that we do not observe. In other words, closing gaps in these measures of academic preparation may not fully close gaps in college enrollment as much the findings would suggest at face value.

We also emphasize that the conditional gaps by race or gender alone do not provide evidence of the presence or absence of discrimination or affirmative action in college admissions. We do not observe every factor that influences college enrollment or admissions. Even controlling for academic and socioeconomic factors, admission is not the only factor that determines college enrollment outcomes; many students who could apply to college do not, and many who are

Across SES, gender, and race, academic GPA alone has the most power to explain enrollment differences

accepted do not attend. In addition, some conditional disparities could be related to students' unobservable characteristics or features of the higher education system that affect groups differently. Racial gaps in academic preparation and SES have been shaped by both structural racism (unequal access to resources and opportunities to learn) and discrimination (unequal treatment based on race) over time. For example, if racial disparities are "explained by" differences in advanced course-taking, we would not conclude that there is no evidence of discrimination or racism, since discrimination also influences access to course-taking and other measures of academic preparation.

College admissions and the cost of college garner considerable attention in public discussions about disparities in college access. While these issues are important, our findings suggest that policymakers and researchers should also pay careful attention to disparities in academic preparation during elementary and secondary education, which are important determinants of college enrollment. Making progress on closing gaps in academic preparation will be particularly important for making progress on closing gender and racial gaps in college enrollment. The same is true for socioeconomic status, though addressing non-academic factors like cost or lack of information will also be important considering that sizable socioeconomic enrollment disparities remain among students with similar academic preparation.

END NOTES

1. Authors' calculations using the 1972 and 2022 Current Population Survey.
2. The Current Population Survey offered one combined category for Asian and Pacific Islander respondents until 2003. For comparability over time, we construct the same category in years after 2003.
3. The HSLS of 2009 is a representative sample of U.S. ninth graders in 2009, not necessarily the graduating class of 2013 or young adults in later years.
4. The choice of 18 months is necessarily arbitrary but captures students who enrolled as "traditional" college students either directly after high school graduation or following a gap year.
5. Students report whether they attend college, which institutions they attended, and when. The HSLS then links the institutions that students report attending to other characteristics about those institutions, for instance the level of that institution (e.g., two-year, four-year). We assign a student to have not enrolled if they said they did not ever attend college by February 2016 (X4EVRATNDCLG), or if they report enrolling in college for the first time after December 2014 (X4PS1START). We assign students to four-year enrollment if they reported ever attending an institution considered to be a four-year institution in the 2015 IPEDS Institutional Characteristics file (S4ICLGLEVEL) before December 2014 and after they graduated high school. We similarly assign students to two-year enrollment, unless they report attending both a two- and four-year institution, in which case we assign them four-year enrollment. Students who do not report attending a two- or four-year institution but do report attending a less than two-year postsecondary institution by December 2014 are included among those who did not enroll in college.
6. Here and throughout the analysis, we use a constant sample of students with non-missing data. We use weights provided by NCES to account for cross-sample attrition. Cross-sectional estimates may vary from full-sample weighted averages due to our sample restriction.
7. We use the HSLS X2SEX variable. If sex is inconsistent across the three sources, X1SEX was assigned based on a review of student names. X2SEX was updated with the follow-up survey in most students' eleventh grade year.
8. We use X4X2SESQ5, which is assigns quintiles based on a revised composite socioeconomic status variable. See HSLS:09 Base-Year to Second Follow-Up Data Documentation for composite construction details.
9. We use X2RACE, which summarizes students' response to a series of dichotomous race questions. Hispanic students include all students who reported that they are Hispanic/Latino/Latina, regardless of race. Every other race/ethnic category includes only non-Hispanic students. The survey uses "Hispanic" in the composite race variable, X2RACE. We follow the terminology used by the HSLS here.
10. We use X3TGPAACAD, X3TGPAMAT, and X3TGPAENG. Deciles are constructing using the corresponding high school transcript weight W3HSTRANS.
11. We use the HSLS X3THIMATH variable, which indicates the highest level of mathematics course taken, to construct students' math course-taking variables. A student is considered to have taken "Advanced math" if the highest math course taken was Trigonometry, Probability and Statistics, or Precalculus. They are considered to have taken "Highly advanced math" if their highest math course was Calculus, an AP/IB math course, or "other advanced math." The variable representing the number of AP or IB Carnegie credits, X3TCREDAPIB, includes fractions of courses. We assign a variable like "one or more AP/IB course" based on if a student had at least one full credit.
12. We convert students' standardized theta math score, X2TXMTSCOR, to a z-score and construct deciles using the W2STUDENT weight, which corresponds to the first follow-up survey when the exam was administered.
13. One exception is that female students have higher college enrollment rates, GPAs, and course-taking, but there is no gender gap in test scores. We return to this below when we discuss the findings by gender.
14. The r-squared from a regression of any enrollment on all the academic preparation measures we consider

here is 0.31, indicating these measures can explain 31% of the variation in that outcome; for four-year enrollment, these measures explain 40% of the variation (not reported).

15. We use a linear regression, even though we have binary outcomes, for ease of interpretation. We follow the recommendation of the survey designer and adjust for clustering in the survey design using the svyset command in STATA. There is not a weight specifically designed for the set of variables we use in this analysis; we use the W4STUDENT weight, which comes close to matching the set of variables we use here. The results are not sensitive to using other reasonable weights.

16. In the discussion, we focus mostly on specifications (1), (2), and (7). We compare across the specification to get a better understanding of which variables are more important in explaining gaps; we interpret these comparisons with some caution, however, since the importance assigned to each variable will depend on the order in which they are added to the regression (Gelbach, 2016).

17. Many researchers take this approach to understand group differences in educational outcomes, including college enrollment or attainment. For example, Goldin et al. (2006) and Jacob (2002) study gender differences; Ellwood and Kane 2000 study differences by SES; Phillips (2011) studies differences by race and SES.

18. See Bohren, Hull, and Imas (2022) and Spriggs (2020) for discussion of why it is difficult to quantify discrimination using this type of regression framework.

19. Recall that we have rescaled the coefficient so that the mean outcome for the omitted category (in this case the third quintile) is equal to the unconditional mean for that group across all of the specifications presented.

20. The results for two-year enrollment (not reported in Figure 3) are difficult to interpret on their own, since some students are on the margin between two-year and no enrollment and some are on the margin between two-year and four-year enrollment. By construction, two-year enrollment is the difference between any and four-year enrollment. The estimates (reported in Table A2) show that rates of two-year enrollment decline with SES, even though rates of any college enrollment increase with SES. Not surprisingly, differences in academic preparation account for a large share—about half—of the difference in two-year enrollment between the fifth and first SES quintiles.

21. We control for the math score in deciles, and there are gender differences in the distribution of scores—males are more likely to be in both the top and bottom decile—that could matter. The overrepresentation of males at the top of the distribution is likely offsetting the overrepresentation at the bottom.

22. Results for all racial groups are available in Table A4.

23. For the regressions presented in Table A4 and Figure 5, we must choose an omitted group, and the coefficients for the other groups indicate the difference in the outcome for that group, relative to the omitted group, conditional on the control variables included in the regression. We chose white to be the omitted category because it is the largest group numerically and do not intend the reader to interpret white students as the “default” or comparisons between whites and other groups as more interesting than comparisons of non-white groups to each other. All groups can be compared to each other and for the figures, we rescale the coefficients by adding the unconditional mean outcome for white students to all the coefficients. The standard errors and significance levels reported with the coefficient in the table are for comparisons to the omitted group (white); to facilitate other comparisons, we report the p-values for each pairwise comparison in Table A4.

24. The magnitude, and in some cases direction, of the conditional disparities does vary somewhat depending on which measures of academic preparation are considered. For example, the gap in favor of Black and Hispanic students is smaller in the specification using only the math test score. See Table A4 for the full results.

25. The interplay between the effects of academic preparation and SES is complex and depends somewhat on which groups are being compared and whether the outcome is any college enrollment or four-year enrollment. For example, controlling for SES in addition to academic preparation matters more for comparisons of Hispanic students to other groups for four-year enrollment than for any enrollment.

REFERENCES

Autor, D., Figlio, D. N., Karbownik, K., Roth, J., & Wasserman, M. (2020). Males at the Tails: How Socioeconomic Status Shapes the Gender Gap (Working Paper No. 27196; Working Paper Series). National Bureau of Economic Research. doi.org/10.3386/w27196

Bailey, M. J., & Dynarski, S. M. (2011). Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. In G. J. Duncan & R. J. Murnane (Eds.), Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances. Russell Sage Foundation. www.nber.org/papers/w17633.

Becker, G. S., Hubbard, W. H. J., & Murphy, K. M. (2010). Explaining the Worldwide Boom in Higher Education of Women. Journal of Human Capital, 4(3), 203–241. https://doi.org/10.1086/657914

Bohren, J. A., Hull, P., & Imas, A. (2022). Systemic Discrimination: Theory and Measurement (Working Paper No. 29820). National Bureau of Economic Research. doi.org/10.3386/w29820.

Carnevale, A. P., Rose, S. J., & Cheah, B. (2011). The College Payoff: Education, Occupations, Lifetime Earnings. The Georgetown University Center on Education and the Workforce. cewgeorgetown.wpenginepowered.com/wp-content/uploads/collegepayoff-completed.pdf

Chetty, R., Hendren, N., Lin, F., Majerovitz, J., & Scuderi, B. (2016). Childhood Environment and Gender Gaps in Adulthood. American Economic Review, 106(5), 282–288. doi.org/10.1257/aer.p20161073

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. doi.org/10.1037/0022-0663.98.1.198

Ellwood, D. T., & Kane, T. J. (2000). Who Is Getting a College Education? In S. Danziger & J. Waldfogel (Eds.), Securing the Future (pp. 283–324). Russell Sage Foundation; JSTOR. www.jstor.org/stable/10.7758/9781610441506.16

Gelbach, J. B. (2016). When Do Covariates Matter? And Which Ones, and How Much? Journal of Labor Economics, 34(2), 509–543. doi.org/10.1086/683668

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

Jacob, B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. Economics of Education Review, 21(6), 589–598. doi.org/10.1016/S0272-7757(01)00051-6

Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The Nonpecuniary Benefits of Schooling. Journal of Economic Perspectives, 25(1), 159–184. doi.org/10.1257/jep.25.1.159.

Owens, J. (2016). Early Childhood Behavior Problems and the Gender Gap in Educational Attainment in the United States. Sociology of Education, 89(3), 236–258. doi.org/10.1177/0038040716650926

Phillips, M. (2011). Ethnic and social class disparities in academic skills. In Diversity in higher education: Toward a more comprehensive approach.

Reardon, S. F., Fahle, E. M., Kalogrides, D., Podolsky, A., & Zárate, R. C. (2019). Gender Achievement Gaps in U.S. School Districts. American Educational Research Journal, 56(6), 2474–2508. doi.org/10.3102/0002831219843824

Reeves, R. V. (2022). Of Boys and Men. Brookings Institution Press.

Spriggs, W. (2020). www.minneapolisfed.org/~/media/assets/people/william-spriggs/spriggs-letter_0609_b.pdf

BROOKINGS

1775 Massachusetts Ave NW,
Washington, DC 20036
(202) 797-6000
www.brookings.edu