

College Admissions Testing and Learning about Ability: Evidence from Strategic ACT and SAT Taking*

Hema Shah[†]

March 31, 2026

Abstract

Using administrative data from North Carolina, I study the decision to take the SAT in the presence of universal ACT testing. I find that low-income students are less likely than their peers to take the SAT in addition to the state-mandated ACT, and that they improve upon their initial ACT scores by less when doing so. Taken together, these disparities decrease low-income students' rankings in the test score distribution when evaluating students on their maximum score rather than their initial ACT score. Income gaps in SAT taking are partially driven by differential responses to the ability signal sent by a student's initial ACT score.

JEL Classification: D83, I23, J24

Keywords: higher education, college admissions testing, college access, ability revelation, learning

*I thank Peter Arcidiacono, Jason Baron, Pat Bayer, Emily Cook, Rob Garlick, Arnaud Maurel, Orgul Ozturk, and Pengpeng Xiao, along with participants of the Duke Labor Lunch, 2023 Association of Policy Analysis and Management annual meeting, and 2023 Southern Economic Association annual meeting, for valuable feedback and suggestions. I thank Kara Bonneau at the North Carolina Education Research Data Center and Tammy Howard, Reenada Perez, and Karl Pond at the North Carolina Department of Public Instruction for data assistance.

[†]Department of Economics, Duke University. 213 Social Sciences Building, Durham, NC 27708. Phone: 757-749-6175. Email: hema.shah@duke.edu.

1 Introduction

Students from low-income backgrounds are less likely to apply to, attend, and graduate from college (Bailey and Dynarski, 2011; Belley and Lochner, 2007; Kena et al., 2015). In the past two decades, the gap in college entry rates between high-income and low-income students has narrowed, yet a difference of over ten percentage points remains between students in the highest and lowest income quintiles (McFarland et al., 2018). The divergence between high-income and low-income students is apparent early in the college decision making process: Even the most high-achieving students from low-income families largely fail to apply to *any* selective postsecondary institution (Hoxby and Turner, 2015). As such, there is an extensive body of literature studying policies designed to improve college access (Page and Scott-Clayton, 2016). A central theme of this literature is the complexity of navigating the college admissions process, particularly for low-income students.

One aspect of the college admissions process that has received significant attention in recent debates is the use of admissions tests. During the Covid-19 pandemic, test-optional policies increased rapidly in prevalence due to disruptions in testing availability (Lovell and Mallinson, 2021). In recent years post-pandemic, colleges and universities have reverted to a spectrum of admissions testing policies (Rosinger et al., 2024). Critics argue that admissions tests, namely the SAT and ACT, disadvantage low-income and minority students in the college admissions process and do not reflect students' true potential to perform well in college coursework. Yet, a long "validity" literature demonstrates that admissions test scores are strongly predictive of college performance (Rothstein, 2004; Westrick et al., 2019; Sacerdote et al., 2025; Chetty et al., 2025; Friedman et al., 2025) and recent work finds no evidence that pandemic-induced test-optional policies benefit low-income students (Sacerdote et al., 2025; Avery et al., 2025).

In this paper, I inform the college admissions testing debate by studying how economically disadvantaged students form their college admissions testing strategies. Using detailed administrative data on North Carolina public school students from 2015-2018, I document selection into taking multiple college admissions tests and the resulting impacts on the dis-

tribution of admissions-relevant¹ test scores. During my sample period, the ACT test was mandated for all 11th grade students in North Carolina public schools. In this setting, I provide the first evidence of strategic multiple test-taking behavior in the presence of universal college admissions testing. Unlike previous research, which studies a selected group of students who choose to take at least one admissions test, I study test-taking behavior among the universe of public high school students in North Carolina. This allows me to compare the distributions of admissions test scores with and without selection into test-taking to shed light on the potentially uneven benefits of universal admissions testing policies.

My analysis begins by showing that economically disadvantaged students are 9 percentage points (30%) less likely than their peers to take the SAT in addition to the in-school ACT, even within high schools and after conditioning on prior academic performance. Conditional on choosing to take multiple tests, economically disadvantaged students improve upon their initial ACT scores by slightly less than their peers. I then analyze how selection into multiple test-taking distorts the admissions-relevant test score distribution in favor of non-disadvantaged students. Theory predicts that more signals should give better quality information about ability, but distortions will arise if only relatively privileged students can take multiple admissions tests. I find that taken together, income gaps in multiple test-taking and score changes move low-income students roughly one percentile lower in the admissions-relevant test score distribution than in the initial ACT score distribution.

I next provide a stylized model of multiple test-taking to generate testable predictions of the mechanisms driving disparities in testing strategy. Then, I test whether disparities in testing strategy are partially driven by students' responses to the ability signal sent by their initial ACT score. Using detailed information on students' lagged academic performance, I predict their scores on the initial in-school ACT test to measure over- and under-performance on the exam. I find that non-disadvantaged male students are more likely to take additional tests when they over-perform on the initial ACT test, but the same is not true for their female and economically disadvantaged peers.

Finally, I perform a simple simulation exercise to show that closing the income gap in

¹I define a student's "admissions-relevant" test score as the maximum of his or her ACT superscore and SAT superscore, borrowing this terminology from [Bloem et al. \(2021\)](#).

multiple test-taking and the resulting score improvements would reduce the income gap in admissions-relevant test scores by 5%. The remainder of the gap can be explained by initial gaps in ACT scores, suggesting that policies designed to increase testing access will have limited effects unless coupled with broader test preparation efforts.

This paper contributes to a small body of literature studying the strategic decision to take multiple college admissions tests (Vigdor and Clotfelter, 2003; Frisancho et al., 2016; Krishna et al., 2018; Goodman et al., 2020; Bloem et al., 2021). I provide novel evidence of multiple test-taking behavior in the presence of universal college admissions testing. The presence of universal ACT testing provides all students with an ability signal and creates a strong default option, which may induce different test-taking patterns among various student groups compared to a setting without universal testing. In particular, because universal testing eliminates positive selection on ability or college-going aspirations, the admissions test-taking population includes many students with limited motivation or resources to take multiple tests. In this setting, disadvantaged students may be less able to retake the exam, leading to lower overall multiple test-taking rates and potentially larger income gaps in test-taking behavior compared to states without universal testing. Indeed, I find that the 45% multiple test-taking rate in my sample is lower than in prior studies. Goodman et al. (2020) find that 54% of all SAT-takers in the U.S. take multiple SAT tests, while Bloem et al. (2021) find that 57% of all admissions test-takers in Georgia take multiple admissions tests.

My results also contribute to the literature on universal college admissions testing policies (Klasik, 2013; Hurwitz et al., 2015; Goodman, 2016; Hyman, 2017; Cook and Turner, 2019; Swiderski, 2025). These policies are intended to expand access to testing and provide information about students' academic preparation for postsecondary education (Hyman, 2017), and studies typically find modest positive effects on college enrollment. I provide a potential explanation for the relatively small and potentially uneven impacts of universal college admissions testing policies on college enrollment: If many successful college applicants take multiple college admissions tests, then the effect of gaining access to a single test may be diluted by selection into multiple test-taking. Nonetheless, substantial benefits remain, as universal ACT testing in North Carolina increased ACT participation from less than 20% to nearly 100% (ACT Research, 2015). Moreover, state reporting shows that economically

disadvantaged students participate in ACT testing at similarly high rates, with roughly one third of disadvantaged students in recent cohorts earning a score above the threshold for admission to state universities ([myFutureNC, 2025](#)).

Finally, this paper adds to the literature on educational decision making and beliefs about ability ([Arcidiacono, 2004](#); [Altonji et al., 2016](#)). To date, this literature has mainly focused on the impacts of beliefs on college major choice and dropout decisions, with analysis of heterogeneity primarily limited to gender differences. Prior work studying the impacts of ability signals generated by standardized tests on students' beliefs and decision making in the K-12 context includes [Papay et al. \(2016\)](#), studying student responses to proficiency labels on standardized mathematics examinations, and [Foote et al. \(2015\)](#), studying the effect of ACT college readiness labels on college enrollment decisions. Most similarly to my study, [Bond et al. \(2018\)](#) estimate the impact of the ability signal sent by a student's SAT score, defining a student's ability signal as the difference between his or her PSAT score and SAT score and finding evidence that positive ability signals lead students to send their SAT scores to more selective portfolios of colleges. I contribute to this literature by providing the first evidence that differential responses to the ability signal sent by a student's initial ACT score contribute to subsequent admissions testing decisions. This result has broad implications for the impacts of student evaluation on educational decision making and later life outcomes. Ultimately, my analysis contributes to our understanding of how universal college admissions tests reveal ability and impact subsequent decisions among students from various backgrounds.

2 Background and Data

2.1 College Admissions Testing in North Carolina

The SAT and ACT tests are standardized exams used to measure United States students' academic preparation in the college application process. The exams are administered by the College Board and the ACT, two private organizations. During my sample period, the SAT and ACT were widely used in selective college admissions. Among U.S. colleges and

universities that are not open enrollment, over 75% of schools either required, recommended, or considered SAT or ACT scores (Bloem et al., 2021).²

Every four-year college and university in the United States accepts the SAT and ACT interchangeably (Goodman, 2016), although the tests differ slightly in content and format. The SAT contains two main sections, Math and Reading, and each section is scored on a scale of 200 to 800 by multiples of 10. Students receive a total score between 400 and 1600 that is the sum of their two main section scores.³ The ACT contains four main sections, English, Math, Reading, and Science, and each section is scored on a scale of 1 to 36. Students receive a composite score between 1 and 36 that is the rounded average of their four main section scores. To compare scores across the two tests, colleges use an SAT-ACT concordance table published by the College Board and the ACT.⁴

Beginning in the 2012-2013 academic year, North Carolina mandated the ACT as part of the state’s school accountability program. The test is now administered in high schools during the school day, free of charge, to all 11th grade students. The test is given during the spring semester of 11th grade, on or near March 1st. For 87% of students in my sample period, which is entirely post-policy change, the first college admissions test they take is the in-school ACT. This is consistent with test timing patterns in other states and the testing companies’ advice that students take their first test in the spring of their junior year of high school (Goodman et al., 2020; Bloem et al., 2021). Hereafter, I use the phrases “in-school ACT” and “initial ACT” interchangeably.

As in all other U.S. states, North Carolina students are allowed to take additional ACT and SAT tests with no limit on the total number of tests taken. In practice, students are limited by the number of exam dates offered per year, which is typically no more than seven for each test. Students may also be constrained by their ability to afford testing fees, as the

²Due to testing disruptions induced by the Covid-19 pandemic, many colleges have relaxed their admissions testing requirements in the years after my sample period.

³Prior to 2016, the SAT contained a third section, Writing, such that total scores ranged from 600 to 2400. In spring 2016, the SAT format was revised to drop Writing and rename Reading “Evidence Based Reading and Writing.” Since this change occurred midway through my sample period, I refer to the two sections as Math and Reading and consider only Math and Reading scores throughout my analysis.

⁴A concordance table corresponding to my sample period can be found here: <https://collegereadiness.collegeboard.org/pdf/guide-2018-act-satconcordance.pdf>. The concordance tables convert SAT scores to ACT scores with the same percentile rank, and vice versa, for a group of students who took both tests.

state of North Carolina pays only for the in-school ACT test and not for additional testing administrations. The cost of additional tests is \$0 for low-income, fee waiver-eligible students and between \$40-\$60 per test for waiver-ineligible students.⁵ Other test-taking costs may include time spent registering for the test, transportation costs, and time taken off from work or extracurricular activities.

While the direct and indirect costs of taking multiple admissions tests are nontrivial, the potential benefits are quite large. Students can improve their scores by maturing or by taking additional coursework between the in-school ACT test and later test administrations. Alternatively, students can improve their scores by gaining familiarity with the standardized admissions testing process. Both the “learning over time” and “learning by doing” explanations have been supported by prior research (Vigdor and Clotfelter, 2003; Frisanco et al., 2016). Furthermore, the most common college admissions testing policy eliminates the risk that a student will decrease his or her test score by taking additional tests. Roughly 75% of colleges and universities follow a policy of “superscoring”, that is, summing a student’s maximum score on each exam section (College Board, 2015). Under superscoring policies, as opposed to policies that consider most recent test scores or average test scores, a student’s admissions-relevant test score can only increase through additional testing attempts. Take, for example, a student who scores a 400 on math and 400 on reading on her first SAT attempt, then on a second attempt scores an 800 on math and 300 on reading. Her SAT superscore would be $800 + 400 = 1200$. Since the most common admissions policy is to superscore within tests but not across the SAT and ACT, a student’s admissions-relevant test score is the maximum of his or her ACT superscore and SAT superscore. This admissions-relevant test score coincides with the maximum composite score from an individual test administration for over 90% of my sample. For brevity, I use the terms “admissions-relevant test score” and “maximum test score” interchangeably.

Admissions test scores are particularly important for North Carolina students planning to attend a public, in-state, 4-year college or university. All public 4-year colleges and universities in North Carolina are part of the University of North Carolina (UNC) system.

⁵SAT fee waivers cover up to two SAT tests, while ACT fee waivers cover up to four ACT tests. Demonstrating free or reduced price lunch (FRL) eligibility is the primary way to obtain SAT and ACT testing fee waivers, although many FRL-eligible students do not take up fee waivers (Goodman et al., 2020).

During my sample period, eligibility for admission to any school within the UNC system required a minimum GPA of 2.5 and a minimum ACT score of 17.⁶⁷ Many public and private universities in North Carolina have admissions standards that are significantly higher than the UNC system-wide threshold; therefore, students stand to benefit from test score increases even above the threshold. For example, the 25th percentile ACT score of entering students at the state flagship university, UNC Chapel Hill, was 28 in 2019. At UNC Charlotte and UNC Wilmington, the 25th percentile ACT score was 22 in the same year.

2.2 Data

I use student-level administrative data on the universe of 11th grade students in North Carolina public schools from 2015-2018, provided by the North Carolina Education Research Data Center (NCERDC).⁸ The data includes each student’s initial ACT score from the in-school ACT administration, highest SAT score from each academic year, state standardized test scores, course-taking and grade histories, demographics, and school characteristics.

The data does not include additional ACT scores from weekend test administrations. Roughly 20% of ACT-takers in North Carolina retake the ACT test during my sample period.⁹ Prior work has shown that, among all U.S. ACT-takers, low-income students retake the ACT test at lower rates than their higher-income peers ([Harmston and Crouse, 2016](#)). Thus, my results should be interpreted as a lower bound on the extent to which multiple test-taking distorts the observed test score distribution in favor of higher-income students. Additionally, the NCERDC data does not include the total number of times a student took

⁶Each college and university in the UNC system is permitted to grant admissions requirement exceptions to up to 1% of admitted students. During my sample period, a pilot program granted additional exceptions to four historically black colleges and universities (HBCUs): Elizabeth City State University, Fayetteville State University, and North Carolina Central University. The pilot program allowed for a sliding scale that offset 10-point SAT score reductions and 1-point ACT score reductions with 0.1-point GPA increases, up to a minimum SAT score of 750 and minimum ACT score of 15. Each of the four universities was permitted to admit up to 100 students per class through the sliding scale.

⁷Students can meet the admissions test score requirement by taking either the SAT or the ACT. The minimum SAT score for admission was 800 until the implementation of the revised SAT test in spring 2016, when the minimum was increased to 880 due to inflation in the revised test scores.

⁸Although universal ACT testing in North Carolina began in 2013, my analysis begins in 2015 because SAT retakes are incompletely recorded in the NCERDC data prior to that year.

⁹Source: Internal calculations provided by ACT. North Carolina’s ACT retake rate is substantially lower than the U.S. average of roughly 45% during the sample period ([Harmston and Crouse, 2016](#)), likely because universal ACT testing reduces positive selection into ACT-taking on motivation and college intent.

the SAT within a particular academic year, only the highest score across all test attempts within a year. Due to these data restrictions, I define multiple test-taking as a binary indicator for taking the SAT at least once in the presence of universal ACT testing and I define “admissions-relevant” or “maximum” test score as the maximum of a student’s initial ACT score and SAT superscore.

I omit from my sample the small percentage of students who are exempted from taking the in-school ACT test.¹⁰ This leaves me with a total sample size of 399,954 students. I convert all ACT scores to the 1600-point SAT scale using the official SAT-ACT concordance tables discussed above.¹¹ This approach preserves information across the full range of test scores, since the finer 1600-point SAT scale reduces the likelihood of tied scores relative to the coarser 36-point ACT scale, allowing for more precise comparisons of student performance.¹²

Table 1 presents sample characteristics by economically disadvantaged status (EDS), defined as eligibility for free or reduced-price lunch. Overall, 45% of students in the sample take multiple admissions tests, but this behavior differs sharply by economically disadvantaged status. 38% of the sample is economically disadvantaged. Among non-disadvantaged students, 55% take multiple tests, compared with only 31% of economically disadvantaged students. On average, economically disadvantaged students have ACT scores that are 142 points lower and admissions-relevant test scores that are 154 points lower than those of non-disadvantaged students, after translating ACT scores to the 1600-point SAT scale. These patterns suggest that disparities in multiple test-taking widen the income gap in test scores.

¹⁰Exemptions to the mandatory ACT policy include students who have a significant cognitive disability and receive instruction in the Extended Content Standards; students who have a current Individualized Education Program (IEP) documenting participation in the Grade 11 College and Career Readiness Alternate Assessment (CCRAA) as well as a written parental request for participation in the CCRAA; students deemed medically fragile because of a significant medical emergency and/or condition and are unable to participate in testing; students who have been retained in the 11th grade and previously took the ACT; and students who took the SAT or the ACT before January 1, 2016, with scores that meet the ACT college readiness benchmark standards. In practice, many students with qualifying prior SAT or ACT scores still opt in to taking the in-school ACT test. My sample of 11th grade ACT-takers includes 95.88% of enrolled 11th grade students without disabilities and 72.13% of students with disabilities.

¹¹The concordance table was revised in spring 2016 when the format of the SAT changed, midway through my sample period. I apply the old concordance table to convert ACT scores from spring 2015 and the new concordance table to later ACT scores.

¹²Appendix Table A3 demonstrates that the main results are robust to converting all scores to the 36-point ACT scale. This suggests that increases in percentile rank from taking the SAT in addition to the ACT are not simply driven by the SAT’s finer 1600-point scale breaking ties between students with the same ACT scores.

Table 1: Sample Characteristics by Economically Disadvantaged Status (EDS)

	Full Sample Mean	Non-Disadvantaged Mean	Disadvantaged Mean
Took Multiple Tests	.4548	.5459	.3066
ACT Score	955.2	1009	867.1
SAT Superscore	1080	1118	969.4
Maximum Test Score	980.8	1039	885.4
Maximum Score - ACT Score	25.63	30.14	18.28
Female	.5061	.5041	.5093
Black	.2468	.158	.3913
Hispanic	.1304	.07677	.2177
White	.5425	.6859	.3091
Asian	.03185	.03583	.02537
Observations	399954	247741	152213

Sample includes first-time 11th grade ACT-Takers in North Carolina public schools from 2015-2018
 ACT and SAT scores are presented on the 1600-point SAT scale

Figure 1 plots kernel density estimates of the distributions of ACT scores, SAT superscores, and admissions-relevant test scores. The SAT score distribution is shifted significantly to the right of the ACT score distribution, suggesting that there is positive selection into taking the SAT. Figure 2 plots kernel density estimates of the test score distributions, restricting the sample to multiple test-takers. Among this selected sample, the ACT score distribution is much closer to the SAT distribution. Still, it is clear that that most students who take multiple tests perform better on the SAT, suggesting that students are improving upon their initial ACT scores by learning over time and learning by doing. On average, multiple test-takers improve upon their initial scores by roughly 56 points on the 1600-point SAT scale. This aligns closely with [Bloem et al. \(2021\)](#), who find that Georgia students who take the ACT once before switching to the SAT improve their initial score by 53 points. These gains are smaller than the 88-point improvement reported by [Goodman et al. \(2020\)](#) and the 66-point improvement reported by [Bloem et al. \(2021\)](#) among SAT-retakers, suggesting that students may achieve larger score increases when retaking the same test due to learning by doing.

It is unlikely that the difference between the SAT score distribution and the ACT score distribution is driven by differences in testing difficulty, as the SAT-ACT score concordance tables are designed to allow comparability across test scores. Therefore, the concordance

tables are set such that each ACT score is converted to an SAT score with the same percentile rank for a group of students who took both tests.

Figure 1: Test Score Distributions, Full Sample

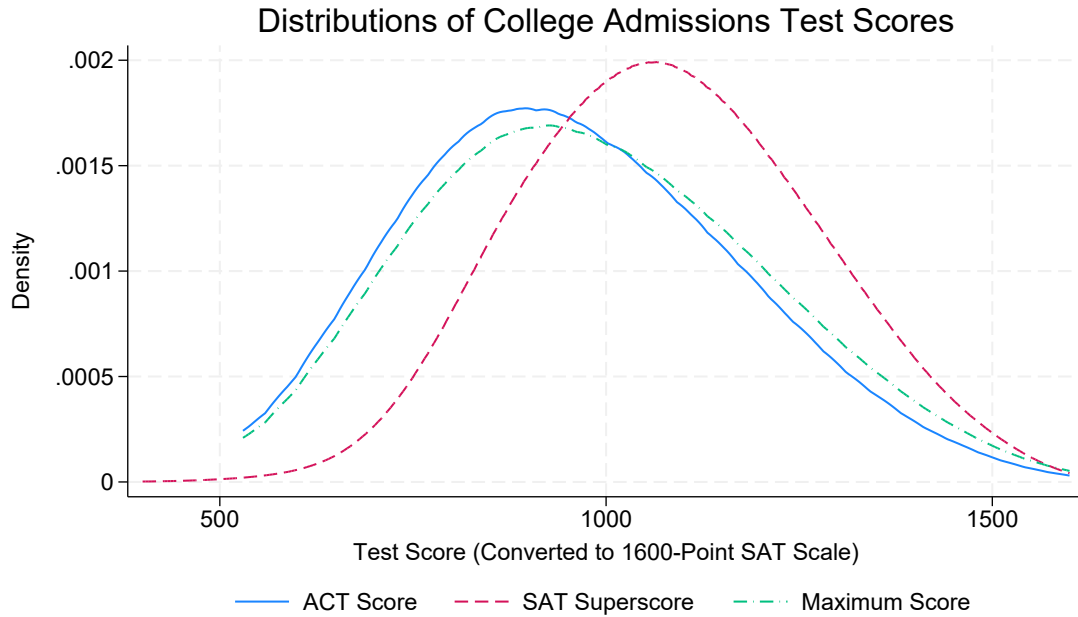
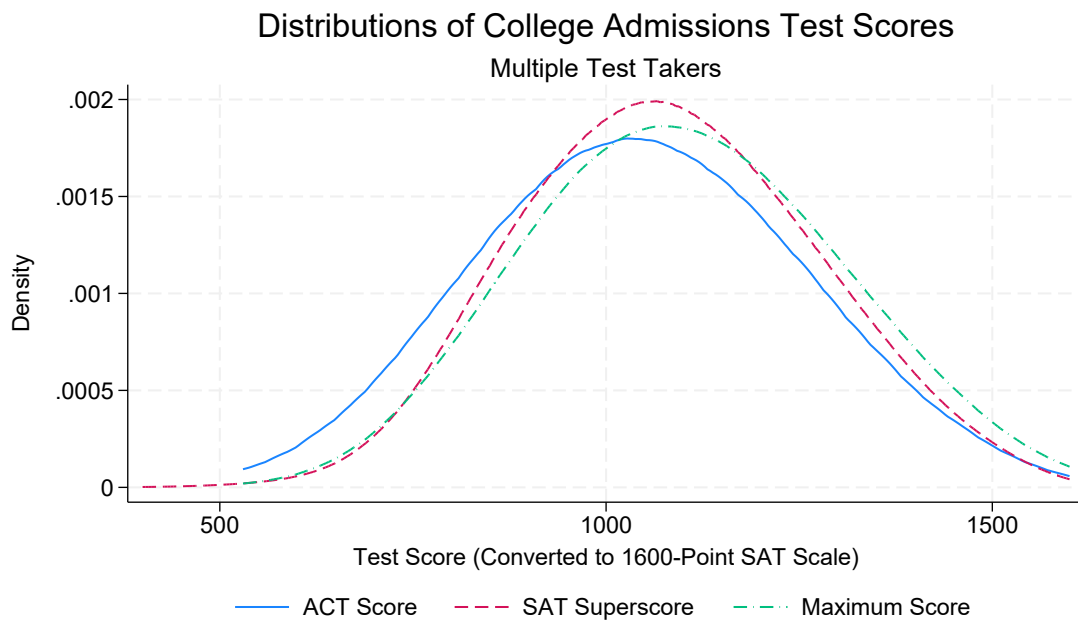


Figure 2: Test Score Distributions, Multiple Test Takers

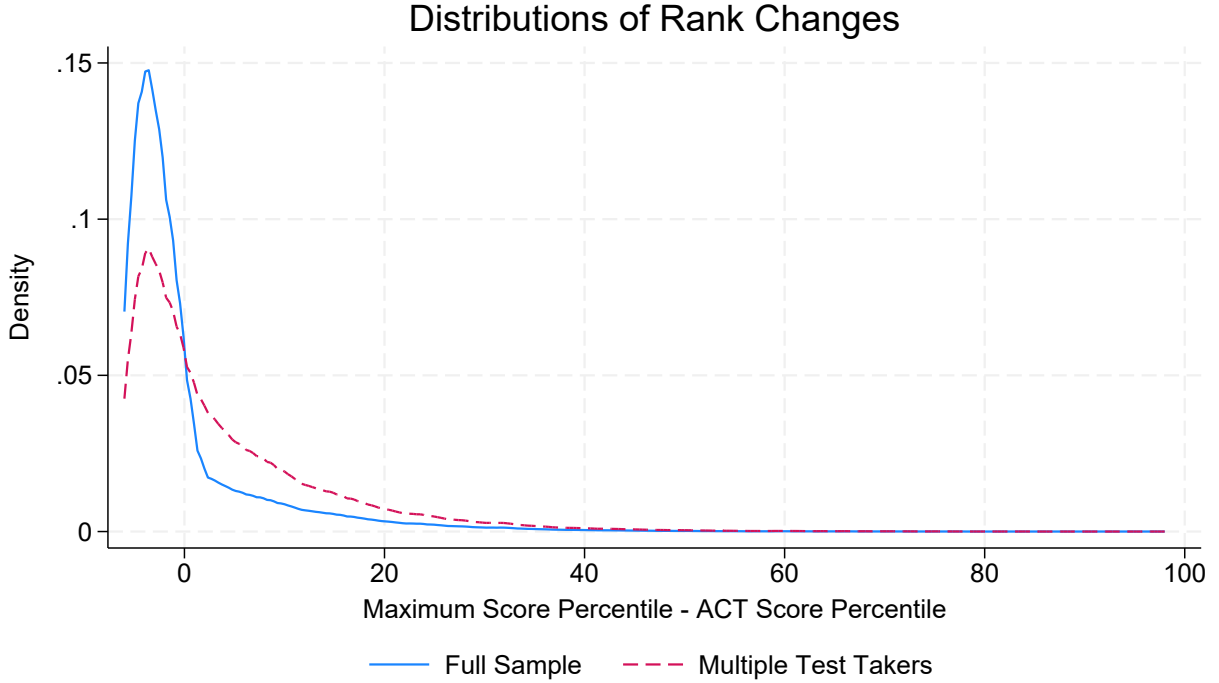


In order to examine distortions in the test score distribution induced by multiple test-taking, I examine changes in test score rank between the admissions-relevant test score and initial ACT score distributions. Specifically, I define a student’s “rank change” as the difference between his or her percentile rank in the admissions-relevant test score distribution and his or her percentile rank in the initial ACT score distribution. A positive value of this variable indicates that a student moved higher in the test score distribution by taking multiple tests, while a negative value indicates that he or she was surpassed by other students who took multiple tests.

Due to the coarse distribution of scores on the 36-point ACT scale, there are inevitably many students tied at the same ACT scores, while the 1600-point SAT scale produces a finer distribution with fewer ties. Because the maximum test score is the higher of a student’s ACT and SAT scores, the inclusion of SAT scores assigns students to more distinct values relative to the ACT score distribution. Under the standard percentile definition, all students with the same score are assigned the percentile corresponding to the top of their tied group. For example, consider a scenario in which 10% of students are tied at the 50th ACT percentile, while only 1% of students are at the 50th percentile of the maximum test score distribution. In the presence of large ACT ties, this convention inflates students’ relative rank within the tied ACT score group and distorts comparisons between initial ACT percentiles and maximum test score percentiles. To address distortions caused by ties in ACT scores, I assign ACT percentiles using the midpoint of each score’s percentile bin rather than the top of the bin. While all students with the same ACT score still receive the same percentile, this midpoint assignment better reflects their relative position within the tied group. As a result, students who improve upon their initial ACT score by taking the SAT move up in the ranking, while those who do not take multiple tests move down relative to their peers.

Figure 3 plots changes in rank in the full sample and among multiple test-takers. Most students move down in rank when taking multiple tests into account, although decreases in rank are all smaller than 6 percentiles. The long right tail of the rank change distribution among multiple test-takers reveals that a small number of students greatly increase their rank in the test score distribution by taking multiple tests. The 99th percentile of the rank change distribution translates to an increase in rank of 33 percentiles.

Figure 3: Distribution of Changes in Test Score Rank



3 Multiple Test-Taking and Rank Distortions

I estimate three descriptive regressions to understand how students’ positions in the test score distribution change when evaluating students using their initial ACT scores versus their admissions-relevant (“maximum”) test scores. Each regression specification considers student i in school s and cohort t . First, I estimate a probit regression of multiple test-taking on a vector of student characteristics, along with school and cohort fixed effects:

$$\mathbb{P}(\text{took multiple tests})_i = \Phi(\beta X_i + \alpha_s + \gamma_t) \tag{1}$$

where X_i includes initial ACT score, 10th grade GPA, gender, race, economically disadvantaged status, and race by disadvantage interactions, α_s are high school fixed effects, and γ_t are cohort fixed effects.

Next, using only the subsample of multiple test-takers, I estimate an OLS regression of score improvement, the difference between a student’s admissions-relevant test score and

initial ACT score, on the same covariates.

$$\text{admissions-relevant score}_i - \text{ACT score}_i = \beta X_i + \alpha_s + \gamma_t + \epsilon_{ist} \quad (2)$$

Taken together, the results of these two regressions inform my analysis of changes in test score rank resulting from multiple test-taking behavior. To evaluate how gaps in multiple test-taking and gaps in score changes combine to distort students' rankings in the admissions-relevant test score distribution, I estimate an OLS regression of change in test score rank on the same covariates.

$$\text{admissions-relevant score rank}_i - \text{ACT score rank}_i = \beta X_i + \alpha_s + \gamma_t + \epsilon_{ist} \quad (3)$$

Coefficient estimates are presented in Table 2. Standard errors are clustered at the school-cohort level in all regressions. Unsurprisingly, students with higher grade point averages are more likely to take multiple tests. Students with higher grades also improve upon their initial ACT scores by more when taking additional tests. Therefore, these students increase in percentile rank, on average, when ranking students on their admissions-relevant test scores instead of their initial ACT scores.

Conditional on prior academic performance, economically disadvantaged students are significantly less likely to take multiple tests than non-disadvantaged students. Translating coefficient estimates to average marginal effects suggests that disadvantaged students are 9.1 percentage points less likely to take multiple tests, roughly 30% relative to the non-disadvantaged multiple test-taking rate. When disadvantaged students do take multiple tests, they improve upon their ACT scores by less than non-disadvantaged students, although the average difference translates to fewer than 7 points on the 1600-point SAT scale. Taken together, these two effects move disadvantaged students 1.2 percentiles lower in the test score distribution, on average, when ranking students on their admissions-relevant test scores instead of their initial ACT scores.

In addition to disparities by disadvantaged status, there are also differences in multiple test-taking behavior by race and gender. Conditional on disadvantaged status and prior academic performance, Black students are significantly more likely than white students to

take multiple tests, while Hispanic students are less likely to do so. Both Black and Hispanic students improve upon their initial ACT scores by less than white students when taking multiple tests, although the differences in score changes are relatively small. Ultimately, multiple test-taking increases test score rank by 1.3 percentiles for Black students and decreases test score rank by 1.1 percentiles for Hispanic students, on average. Asian students are more likely than white students to take multiple tests, and on average they improve upon their initial scores by approximately 23 SAT points (or roughly half of an ACT point) more than their peers. Therefore, multiple test-taking increases test score rank by 1.8 percentiles for Asian students. These patterns do not hold for economically disadvantaged Asian students, who make up less than one third of Asian students and less than one percent of the full sample.

Female students are more likely than male students to take multiple tests, although the gender gap in multiple test-taking is much smaller than disparities by disadvantaged status and race. The gender gap in score improvements, on the other hand, is larger than disparities by disadvantaged status and race. Female students improve upon their initial ACT scores by approximately 20 SAT points (half of an ACT point) less than male students, on average, when taking multiple tests. This disparity decreases test score rank for female students by 0.8 percentiles, on average, when ranking students on their admissions-relevant test score instead of their initial ACT scores. This gender gap in rank change is conditional on prior academic performance, which is higher for female students on average.

Appendix A shows that the results are robust to the inclusion of polynomials in ACT score and GPA and to using a linear probability model instead of the probit model in specification (1).

Table 2: Multiple Test-Taking and Student Characteristics

	(1)	(2)	(3)
	Took Multiple Tests	Score Change	Rank Change
ACT Score	0.00150*** (0.0000264)	-0.219*** (0.00206)	-0.0152*** (0.000225)
10th Grade GPA	0.634*** (0.00574)	25.26*** (0.382)	2.707*** (0.0381)
Female	0.145*** (0.00526)	-20.21*** (0.360)	-0.799*** (0.0295)
Black	0.424*** (0.0118)	-5.328*** (0.614)	1.315*** (0.0668)
Economically disadvantaged (EDS)	-0.422*** (0.00919)	-6.921*** (0.563)	-1.232*** (0.0379)
Black \times EDS	0.247*** (0.0133)	1.863* (0.885)	0.195* (0.0781)
Hispanic	-0.208*** (0.0137)	-5.951*** (0.819)	-1.050*** (0.0671)
Hispanic \times EDS	0.278*** (0.0172)	-1.293 (1.118)	0.508*** (0.0736)
Asian	0.231*** (0.0214)	22.61*** (1.662)	1.762*** (0.123)
Asian \times EDS	0.169*** (0.0366)	-18.03*** (2.390)	-1.012*** (0.203)
Model	Probit	OLS	OLS
N	380465	177273	381599
R^2 /Pseudo- R^2	0.286	0.222	0.132

Coefficients from probit and OLS regressions corresponding to equations (1), (2), and (3)

All models include school and cohort fixed effects

Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The previous results demonstrate that, on average, disadvantaged students are less likely than their peers to take multiple tests, and that they improve upon their initial scores by less when they do take multiple tests. Taken together, these two effects move the average disadvantaged student down in percentile rank when considering multiple test scores. Whether such test score and rank changes matter depends on how test scores are used in admissions. When a test is used as a strict eligibility screen, such as the UNC system-wide requirement that applicants score at least 17 on the ACT, what matters is the score itself, not a student's percentile rank. In contrast, at selective campuses such as UNC Chapel Hill that must ration a limited number of seats among many applicants who all clear this minimum, admissions effectively become rank-based. Students are admitted from the top of the score distribution until seats are filled, so relative standing among applicants matters.

Changes in test scores and rank are therefore only economically relevant for students whose initial scores place them near either a score-based cutoff or a rank-based competitive margin. Students far below any cutoff are unaffected by small changes in score or rank, while students far above the competitive range are likewise unlikely to be affected. The average ACT percentile of disadvantaged students is 39, which corresponds to an ACT score of approximately 16. This suggests that the average disadvantaged student lies close to the UNC system ACT threshold, such that small score improvements from multiple test-taking can determine whether the threshold is crossed. Conditional on crossing that threshold, the same differences may also affect a student's relative standing among other eligible applicants at selective campuses.

While the NCERDC data does not include any details on the specific colleges students plan to attend, I observe a measure of intended 4-year college enrollment from a graduation survey administered during the spring of 12th grade. Descriptive evidence indicates that 4-year college enrollment is substantially higher among students who take multiple college admissions tests: 76% of students who take multiple tests intend to enroll in a 4-year college compared to 23% of students who take only the in-school ACT. Moreover, Appendix B demonstrates that the income gap in 4-year college enrollment is slightly smaller among multiple test-takers when comparing students with the same initial ACT scores. Higher college enrollment rates and smaller income gaps in college enrollment among multiple test-

takers could be due to positive selection on motivation and college intent or to gains in test scores and test score rank from multiple test-taking. Appendix B further plots income gaps in intended 4-year college enrollment by initial ACT score among multiple test-takers who do not improve upon their initial score. Among these students who select into multiple test-taking but do not experience test score or rank gains from retesting, the income gap in 4-year college enrollment is similar in size to the income gap among the full sample of multiple test-takers. This pattern suggests that the smaller income gap in 4-year college enrollment observed among multiple test-takers, relative to single test-takers, is primarily driven by selection into multiple test-taking. Nonetheless, test score and rank gains from multiple test-taking likely contribute to income gaps in college enrollment on the margin.

In the absence of detailed college enrollment data, I further evaluate the extent to which changes in rank due to multiple test-taking are occurring at margins relevant to college enrollment and selectivity by estimating an OLS regression of rank change on the interactions between disadvantaged status and a set of dummy variables indicating initial ACT score decile. Coefficients are relative to the lowest ACT score decile, in which the average rank change is approximately 1.6 percentiles for disadvantaged students and 2.6 for non-disadvantaged students. Changes in rank due to multiple test-taking reflect two types of movements: increases in rank due to one’s own score improving and decreases in rank due to peers’ scores improving while one’s own score stays the same. To disentangle these two types of movements, I estimate equation (4) first among the full sample and then separately among multiple test-takers and single test-takers.

$$\begin{aligned}
& \text{admissions-relevant score rank}_i - \text{ACT score rank}_i \\
& = \beta_{\text{ACT}} \text{ACT score decile}_i + \beta_{\text{EDS}} \text{EDS}_i + \beta_{\text{interact}} \text{ACT score decile}_i \times \text{EDS}_i \\
& + \beta_X X_i + \alpha_s + \gamma_t + \epsilon_{ist}
\end{aligned} \tag{4}$$

Here, X_i includes 10th grade GPA, gender, race, and race by disadvantage interactions. α_s are high school fixed effects, and γ_t are cohort fixed effects. Coefficient estimates of β_{ACT} , β_{EDS} , and β_{interact} are presented in Figures 4 and 5, respectively.

Figure 4 demonstrates that rank change is decreasing in initial ACT decile. This is

unsurprising, given that students who perform very well on the initial ACT test likely have more difficulty improving upon their initial scores. Rank change among multiple test-takers is strongly decreasing in initial ACT decile, underscoring the mechanical inability of students with high initial ACT scores to significantly increase their scores on subsequent attempts. While all single test-takers mechanically decrease in rank when considering multiple test scores, students at the middle of the initial ACT score distribution see the largest decreases.

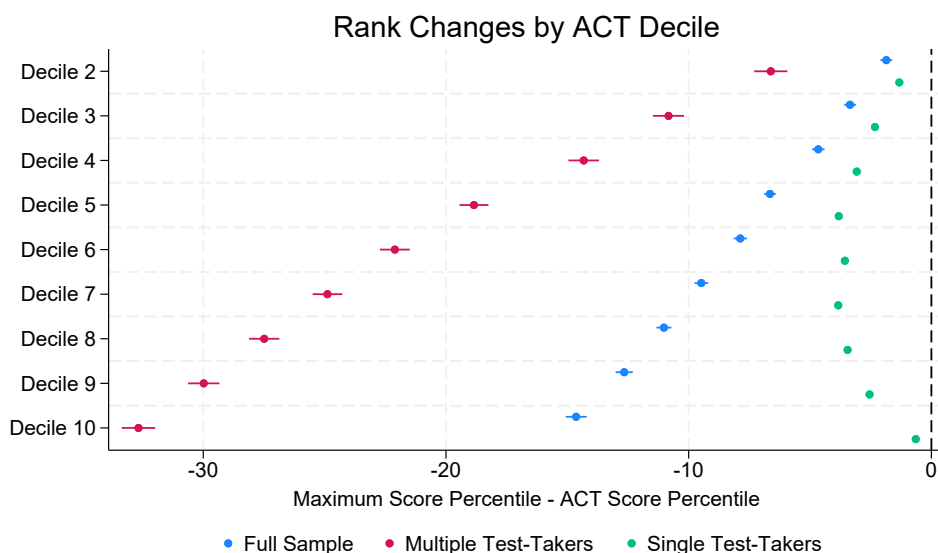
Figure 5 demonstrates that disparities in rank change by disadvantaged status are highly nonlinear in initial ACT decile. Even after controlling for GPA, gender, race, and school and cohort fixed effects, disadvantaged students in deciles 3 through 7 of the initial ACT score distribution decrease in rank relative to their peers due to multiple test-taking.

Disadvantaged multiple test-takers with high initial ACT scores increase in rank relative to non-disadvantaged multiple test-takers when considering multiple test scores. This suggests that, although disadvantaged students are less likely than their peers to take multiple tests, those who do so benefit from subsequent attempts. This could be driven by positive selection of disadvantaged students into multiple test-taking on unobservables such as familial educational background. This selection explanation suggests that not all disadvantaged students would increase in rank, relative to non-disadvantaged students, from multiple test-taking. Alternatively, relative rank increases could reflect larger gains to disadvantaged students from multiple test attempts, perhaps due to low quality in-school ACT testing conditions in schools with more disadvantaged students. This explanation suggests that disadvantaged students who do not take multiple tests, particularly those with high initial ACT scores, would benefit from doing so. Appendix C demonstrates that average disparities in score changes by disadvantaged status are small within initial ACT scores. This result supports the selection explanation because it suggests that the large rank gains for disadvantaged multiple test-takers arise from the ability levels of students who choose to retest, rather than from systematically larger gains among disadvantaged students.

Interaction coefficients for single test-takers are small, suggesting that disadvantaged and non-disadvantaged single test-takers decrease in rank similarly when considering multiple test scores. This result supports the previous evidence that disparities in rank change by disadvantaged status are driven by disparities in multiple test-taking rather than disparities

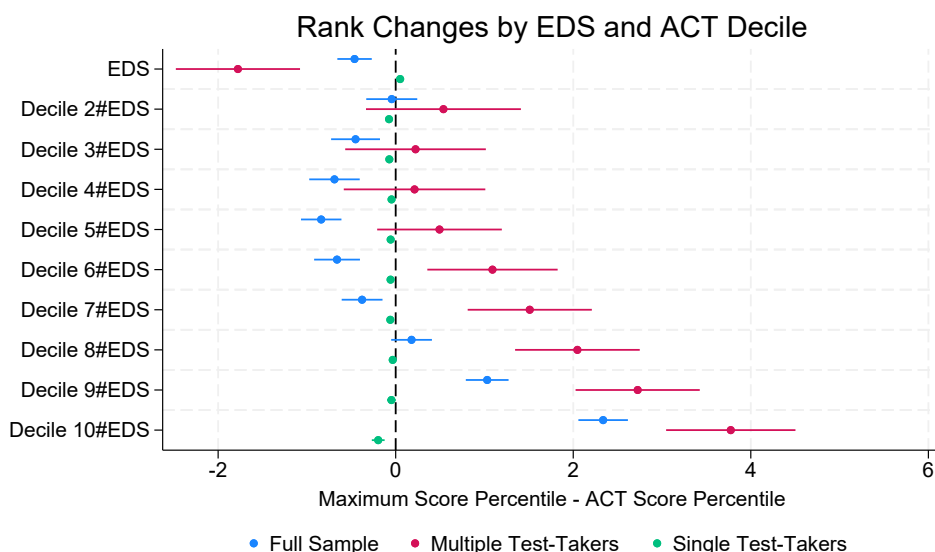
within the multiple test-taking and single test-taking groups.

Figure 4: Coefficients on Initial ACT Decile



Notes: Coefficients from OLS regressions corresponding to equation (4). Excluded category is ACT decile 1, corresponding to students with the lowest initial ACT performance. Outcome is the difference between maximum test score percentile and initial ACT score percentile.

Figure 5: Coefficients on EDS and EDS \times Initial ACT Decile



Notes: Coefficients from OLS regressions corresponding to equation (4). Excluded category is ACT decile 1, corresponding to students with the lowest initial ACT performance. Outcome is the difference between maximum test score percentile and initial ACT score percentile.

4 A Stylized Model of Multiple Test-Taking

I provide a stylized model of multiple test-taking behavior building on [Vigdor and Clotfelter \(2003\)](#). Suppose student i has true academic ability ρ_i , which I assume to be one-dimensional for simplicity. College admissions test scores are noisy estimates of ρ_i , which is unknown to the student. Each time student i takes a college admissions test, she receives a score p_i drawn from a distribution $f_i(p)$ with mean ρ_i . After taking the mandatory in-school ACT test and receiving score p_i^* , student i decides to take an additional admissions test at time t if and only if the expected benefit exceeds the test-taking cost c_i ; that is, if

$$\tilde{\mathbb{E}}_{it}[V_i(a(\max\{p, p_i^*\}) - a(p_i^*)) \mid \mathcal{F}_{it}] > c_i \quad (5)$$

The expected benefit of taking an additional admissions test is equal to student i 's valuation of college admission, V_i , times student i 's expected change in admissions probability from taking an additional admissions test. The student's expected change in admissions probability depends on two unknown objects: $a(\cdot)$, admissions probability as a function of admissions test score, and p , the score received on the additional admissions test. Student i forms subjective expectation $\tilde{\mathbb{E}}_{it}$ over $a(\cdot)$ and p using time t information \mathcal{F}_{it} . \mathcal{F}_{it} contains information about the college admissions process as well as information about student i 's academic ability. Information about the college admissions process is unobserved in my data and can come from family, friends, teachers, counselors, and other sources. Information about academic ability consists of observable pre-ACT signals from standardized test scores and course grades as well as an observable signal from the in-school ACT test itself. Information about academic ability may also include signals, such as SAT tutoring participation, that are unobserved in my data. This model suggests four potential mechanisms driving income gaps in multiple test-taking.

First, direct test-taking costs c_i may vary with disadvantaged status for multiple reasons. Economically disadvantaged students who do not register for testing fee waivers may face budget constraints when paying the testing fee. Economically disadvantaged students may also have difficulties securing transportation to testing centers or face higher opportunity costs of test-taking due to part-time jobs or family responsibilities. Second, economically

disadvantaged students may differ in their valuation of college V_i .

Third, beliefs about admissions probability function $a(\cdot)$ may vary with disadvantaged status if economically disadvantaged students have access to less information about the college admissions process. This could be the case, for example, if college counseling quality is lower in schools with more disadvantaged students.

The fourth mechanism through which disparities in multiple test-taking may arise is through the belief updating channel. Specifically, economically disadvantaged students may update their subjective expected test scores p differently in response to their initial ACT scores. Since admissions probability $a(\cdot)$ is weakly increasing over the support of test scores p , the expected benefit of taking an additional admissions test is increasing in student i 's subjective expected test score at the time of the test-taking decision, $\tilde{\mathbb{E}}_{it}[p | \mathcal{F}_{it}]$. I assume students process information in a Bayesian manner, so that $\tilde{\mathbb{E}}_{it}[p | \mathcal{F}_{it}]$ is a weighted average of pre-ACT and post-ACT beliefs, as follows:

$$\tilde{\mathbb{E}}_{it}[p | \mathcal{F}_{it}] = \alpha_{i,\text{pre-ACT}} \tilde{\mathbb{E}}_i[p | \mathcal{F}_{i,\text{pre-ACT}}] + \alpha_{i,\text{ACT}} \tilde{\mathbb{E}}_i[p | \mathcal{F}_{i,\text{ACT}}] \quad (6)$$

$$\alpha_{i,\text{pre-ACT}} = \frac{\text{var}(p | \mathcal{F}_{i,\text{ACT}})}{\text{var}(p | \mathcal{F}_{i,\text{ACT}}) + \text{var}(p | \mathcal{F}_{i,\text{pre-ACT}})} \quad (7)$$

$$\alpha_{i,\text{ACT}} = \frac{\text{var}(p | \mathcal{F}_{i,\text{pre-ACT}})}{\text{var}(p | \mathcal{F}_{i,\text{ACT}}) + \text{var}(p | \mathcal{F}_{i,\text{pre-ACT}})} \quad (8)$$

Disparities in belief updating by disadvantaged status could arise for two reasons. First, disadvantaged and non-disadvantaged students may have pre-ACT beliefs $\tilde{\mathbb{E}}_i[p | \mathcal{F}_{i,\text{pre-ACT}}]$ that are biased differently. For example, disadvantaged students may be overly confident in their academic abilities if schools with more disadvantaged have less rigorous coursework.

Second, the variance of pre-ACT ability information may vary with disadvantaged status. Since students are Bayesian updaters, $\tilde{\mathbb{E}}_{it}[p | \mathcal{F}_{it}]$, and consequently the decision to take an additional college admissions test, will place more weight on the signal sent by the ACT if pre-ACT information is more diffuse. If standardized test scores and course grades are less predictive of college readiness for disadvantaged students, then disparities in belief updating may arise. Although I am unable to disentangle biased beliefs from diffuse priors, I test whether disadvantaged students respond differently to initial ACT performance.

5 Belief Updating

Restricting my analysis to the subsample of students whose first admissions test taken is the in-school ACT, I examine multiple test-taking behavior in response to the ability signal sent by students' initial ACT scores. Specifically, I test whether disadvantaged and non-disadvantaged students respond differently to underperformance or overperformance on the initial ACT test. To measure the ability signal sent by a student's initial ACT score, I construct a predicted ACT score measure based on past academic performance. I predict ACT performance using an OLS regression of ACT score on cubic polynomials in lagged GPA and test scores, with high school and cohort fixed effects.

$$\begin{aligned}
 \widehat{ACT}_i = & \sum_{k=1}^3 \hat{\beta}_{GPA9,k} GPA9_i^k + \sum_{k=1}^3 \hat{\beta}_{GPA10,k} GPA10_i^k \\
 & + \sum_{k=1}^3 \hat{\beta}_{MA,k} MA_i^k + \sum_{k=1}^3 \hat{\beta}_{RD,k} RD_i^k + \sum_{k=1}^3 \hat{\beta}_{BIOL,k} BIOL_i^k \\
 & + \sum_{k=1}^3 \hat{\beta}_{ALG,k} ALG_i^k + \sum_{k=1}^3 \hat{\beta}_{ENG,k} ENG_i^k \\
 & + \hat{\beta}_X X_i + \hat{\alpha}_s + \hat{\gamma}_t
 \end{aligned} \tag{9}$$

Here, $GPA9_i$ is 9th grade GPA, $GPA10_i$ is 10th grade GPA, MA_i refers to 8th grade math test score, RD_i refers to 8th grade reading test score, $BIOL_i$ refers to 9th grade biology test score, ALG_i refers to 9th grade algebra test score,¹³ and ENG_i refers to 10th grade English test score. X_i includes gender, race, economically disadvantaged status, and race by disadvantage interactions. Appendix E provides more detail on the fit of the prediction model, which predicts the ACT score distribution well with an R^2 of 0.782 and has similar fit between disadvantaged and non-disadvantaged students. A positive value of the residual indicates that a student overperformed on the ACT test relative to her past academic

¹³Prior to the 2012-2013 school year, Algebra 1 was the standard mathematics course taken in 9th grade, or as early as 7th grade for students on advanced math tracks. In the 2012-2013 school year, North Carolina implemented a revised math curriculum and renamed the 9th grade math course "Math 1." The first cohort in my sample to take the revised 9th grade math curriculum is the 2016 cohort. I account for cross-year differences in course curriculum by normalizing all standardized test scores to have mean zero and standard deviation one within each cohort.

performance.

To understand the relationship between multiple test-taking and ACT performance relative to expected performance, I estimate a probit regression of multiple test-taking on predicted ACT score, ACT score residual, and the interaction between the residual and disadvantaged status, shown in equation (10). I estimate the regression on the subsample of students with predicted ACT scores in the middle 50% of the predicted score distribution, as students with low academic achievement are unlikely to attend college and therefore unlikely to take multiple admissions tests and students with very high academic achievement may plan to take multiple admissions tests regardless of their initial performance to be competitive for selective college admissions. Results are presented in Table 3. I estimate the regression separately on the female and male samples. Coefficient estimates suggest that the response to the ACT score residual is positive and larger for males; that is, students who overperform on the initial ACT test are more likely to take subsequent admissions tests. Among non-disadvantaged males, converting coefficient estimates to average marginal effects suggests that a one standard deviation increase in ACT score residual increases multiple test-taking by roughly 2 percentage points. Among disadvantaged males and both disadvantaged and non-disadvantaged females, the effect size is halved.¹⁴

$$\mathbb{P}(\text{took multiple tests})_i = \Phi(\beta_{predict} \hat{ACT}_i + \beta_{resid} \hat{\epsilon}_{ist} + \beta_{EDS} EDS_i + \beta_{interact} \hat{\epsilon}_{ist} \times EDS_i + \alpha_s + \gamma_t) \quad (10)$$

To further understand how multiple test-taking behavior varies across the distribution of ACT score residuals, I estimate a probit regression of multiple test-taking on predicted ACT score and the interactions between disadvantaged status and a set of dummy variables indicating ACT score residual decile, shown in equation (11). Coefficients are relative to the lowest decile of the ACT score residual distribution, in which the multiple test-taking rate is approximately 31% for disadvantaged students and 47% for non-disadvantaged students.

¹⁴Appendix D demonstrates that the results are robust to using a linear probability model and to allowing each predictor of initial ACT score to enter separately instead of summarizing them in the predicted ACT score measure. Results are qualitatively similar but smaller in magnitude among the full sample.

Table 3: Multiple Test-Taking and ACT Score Residual

	(1)	(2)
	Female	Male
Predicted ACT Score	0.144*** (0.00308)	0.144*** (0.00314)
ACT Score Residual	0.0433*** (0.00307)	0.0644*** (0.00266)
Economically disadvantaged	-0.207*** (0.0110)	-0.204*** (0.0117)
Economically disadvantaged \times Residual	-0.0136** (0.00465)	-0.0160*** (0.00414)
Model	Probit	Probit
N	79766	76811
Pseudo- R^2	0.122	0.124

Coefficients from Probit regressions corresponding to equation (10)

All models include school and cohort fixed effects

Sample includes students in the 25th-75th percentiles of predicted ACT scores

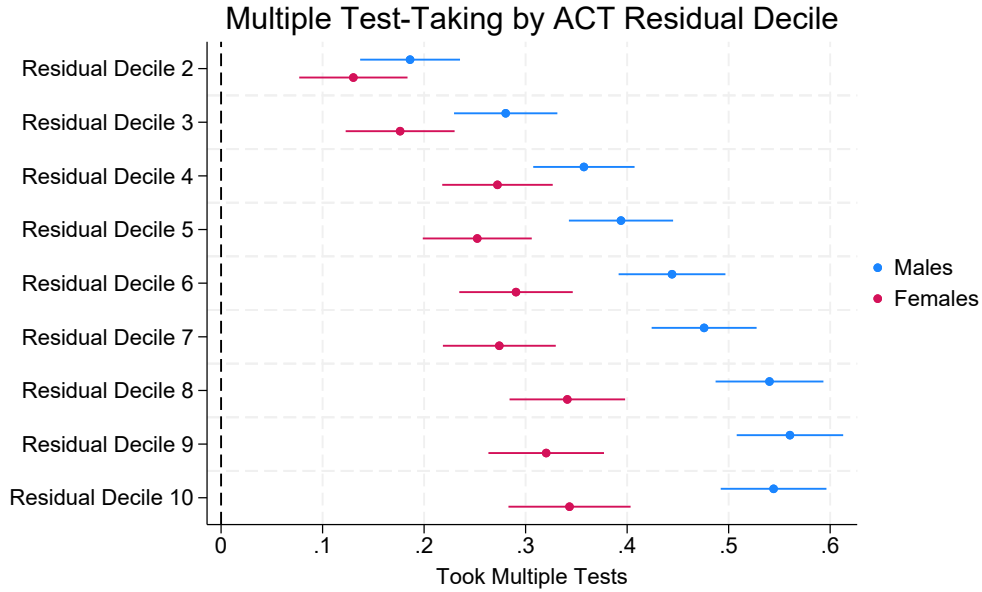
Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

$$\begin{aligned} \mathbb{P}(\text{took multiple tests})_i = & \Phi(\beta_{predict} \hat{ACT}_i + \beta_{resid} \text{residual decile}_i + \beta_{EDS} EDS_i \\ & + \beta_{interact} \text{residual decile}_i \times EDS_i + \alpha_s + \gamma_t) \end{aligned} \quad (11)$$

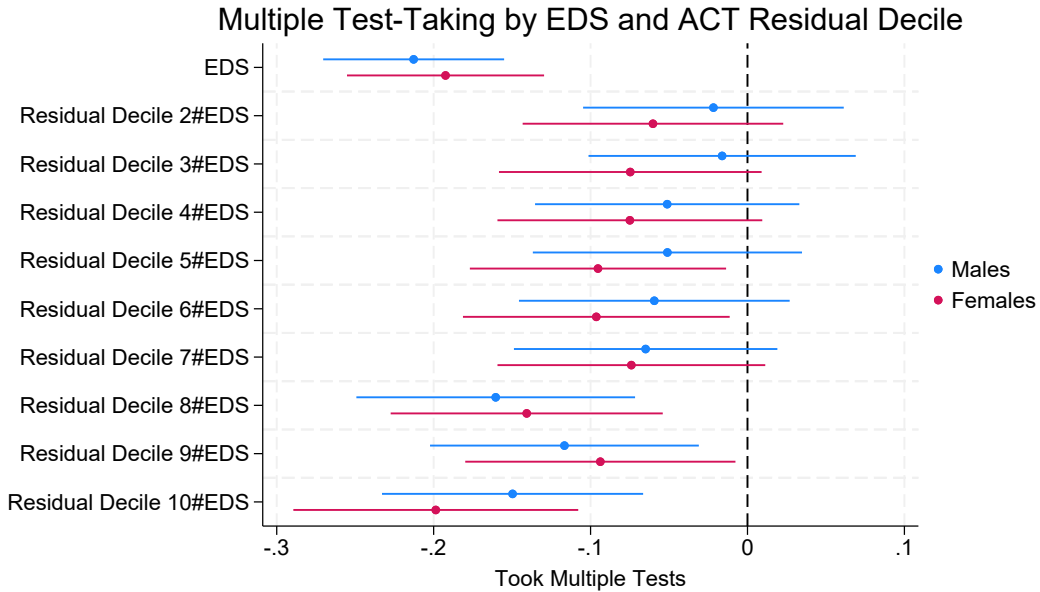
Estimates of β_{resid} , β_{EDS} , and $\beta_{interact}$ are presented in Figures 6 and 7. Estimates of β_{resid} suggest that female students' multiple test-taking behavior varies nonlinearly with ACT score residual. Female students in the lowest residual deciles have the lowest multiple test-taking rates conditional on predicted ACT score; that is, female students who perform very poorly on the initial ACT test relative to expectations are less likely than their peers to take additional tests. Among male students, the multiple test-taking rate conditional on predicted ACT score is more strongly increasing in ACT score residual; that is, male students who perform well on the initial ACT test relative to expectations are more likely than their peers to take additional tests, even among those who overperform significantly. Estimates of $\beta_{interact}$ suggest that the increasing response to ACT residual is stronger among non-disadvantaged students.

Figure 6: Coefficients on Residual Decile



Notes: Coefficients from probit regressions corresponding to equation (11). Excluded category is residual decile 1, corresponding to students who underperformed relative to their past academic performance. Outcome is an indicator for multiple test-taking.

Figure 7: Coefficients on EDS and EDS \times Residual Decile



Notes: Coefficients from probit regressions corresponding to equation (11). Excluded category is residual decile 1, corresponding to students who underperformed relative to their past academic performance. Outcome is an indicator for multiple test-taking.

6 Closing the Income Gap in Multiple Test-Taking

How much would the income gap in admissions-relevant test scores close if disadvantaged students took multiple admissions tests at the same rate as observably similar non-disadvantaged students? To understand this question, I perform a simple simulation exercise. I equate multiple test-taking rates by disadvantaged status among students in the same cohort with the same initial ACT score as well as the same GPA, gender, and race. To understand how initial ACT score, GPA, gender, and race impact multiple test-taking rates among non-disadvantaged students in the true data, I estimate a probit regression of multiple test-taking on initial ACT score, 10th grade GPA, gender, race, and cohort fixed effects, estimated on the sample of non-disadvantaged students only. I compare students across schools rather than including school fixed effects; therefore, this simulation exercise can be partially interpreted as closing cross-school gaps in multiple test-taking behavior.

$$\mathbb{P}(\text{took multiple tests})_i = \Phi(\beta_{ACT}ACT_i + \beta_{GPA}GPA_i + \beta_X X_i + \gamma_t) \quad (12)$$

Here, X_i includes race and gender. I simulate multiple test-taking among both disadvantaged and non-disadvantaged students using the estimated coefficients from this regression, $\hat{\beta}$ and $\hat{\gamma}$, and students' true characteristics taken from the data. This equates multiple test-taking rates of disadvantaged students with those of non-disadvantaged students with the same observable characteristics. Specifically, I simulate multiple test-taking for student i in cohort t using a Bernoulli random variable with mean $\hat{\pi}$ defined as follows.

$$\hat{\pi}_i = \Phi(\hat{\beta}_{ACT}ACT_i + \hat{\beta}_{GPA}GPA_i + \hat{\beta}_X X_i + \hat{\gamma}_t) \quad (13)$$

I simulate score improvements conditional on multiple test-taking analogously, equating mean score improvements by disadvantaged status among students in the same cohort with the same initial ACT score as well as the same GPA, gender, and race. To understand how score improvements vary with initial ACT score, GPA, gender, and race among non-disadvantaged students in the true data, I estimate an OLS regression of score improvement on initial ACT score, 10th grade GPA, gender, race, and cohort fixed effects (but not school

fixed effects), estimated on the sample of non-disadvantaged students only.

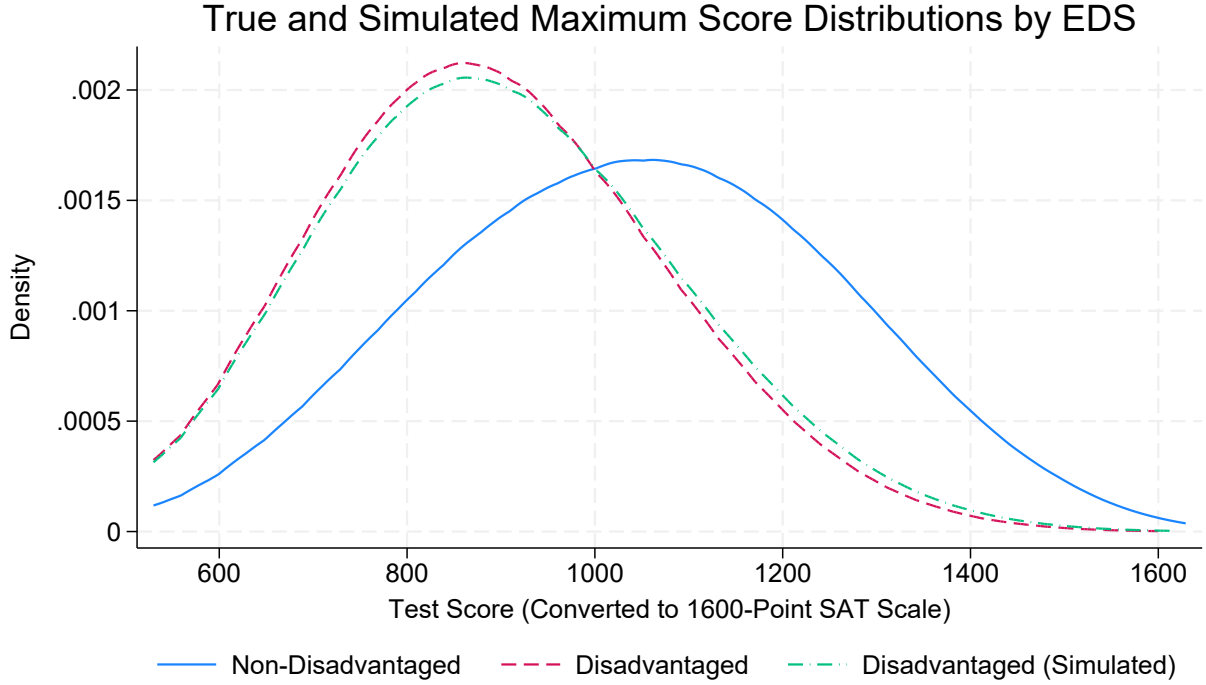
$$\text{score change}_i = \delta_{ACT}ACT_i + \delta_{GPA}GPA_i + \delta_X X_i + \xi_t + \epsilon_{it} \quad (14)$$

X_i includes race and gender. I simulate score improvements, conditional on multiple test-taking, among both disadvantaged and non-disadvantaged students using the estimated coefficients from this regression, $\hat{\delta}$ and $\hat{\xi}$, and students' true characteristics from the data.

$$\widehat{\text{score change}}_i = \hat{\delta}_{ACT}ACT_i + \hat{\delta}_{GPA}GPA_i + \hat{\delta}_X X_i + \hat{\xi}_t \quad (15)$$

Figure 8 plots kernel density estimates of the true and simulated test score distributions by disadvantaged status. The simulation shifts the admissions-relevant test score distribution among disadvantaged students to the right. Table 4 demonstrates that equating multiple test-taking rates and score changes by disadvantaged status, conditional on GPA, gender, and race, closes the gap in multiple test-taking rates by approximately 9.8 percentage points (41%), from 24.2 percentage points to 14.4 percentage points, and closes the income gap in admissions-relevant test scores by 10 SAT points (7%). The remainder of the gap in multiple test-taking is due to economically disadvantaged students having lower grade point averages, which is negatively correlated with multiple test-taking rates. Similarly, the remaining gap in admissions-relevant test scores can be explained by differences in initial ACT scores.

Figure 8: Simulated and Observed Test Score Distributions



Notes: Simulated maximum scores calculated by simulating multiple test-taking using equation (13) and score change conditional on multiple test-taking using equation (15)

Table 4: Simulated Multiple Test-Taking by EDS

	Non-Disadvantaged Mean	Disadvantaged Mean	Difference
Observed			
Took Multiple Tests	0.5548	0.3129	0.2419
Score Change	61.61	66.91	-5.300
Maximum Test Score	1043	888.2	154.8
Simulated			
Took Multiple Tests	0.5541	0.4104	0.1438
Score Change	61.61	73.88	-12.28
Maximum Test Score	1043	898.3	144.7
Observations	247700	152200	95530

Simulated multiple test-taking indicator calculated using equation (13)

Simulated score change calculated using equation (15)

Simulated maximum score calculated by adding initial ACT score to the product of simulated score change and simulated multiple test-taking indicator

7 Conclusion

Using administrative data on the universe of public school students in North Carolina, I provide the first evidence of income disparities in strategic multiple test-taking behavior under universal ACT testing. I find that low-income students are less likely than their peers to take the SAT in addition to the state-mandated ACT, and that they improve upon their initial ACT scores by less when doing so. These disparities persist within schools and after conditioning on prior academic achievement; as a result, the widely used admissions practice of superscoring moves economically disadvantaged students down in test score rank relative to observably similar non-disadvantaged students.

Additionally, I show that disparities in multiple test-taking behavior by gender and disadvantaged status are partially driven by differential responses to the ability signal sent by a student's initial ACT score. I leverage detailed information on students' past academic performance to predict their scores on the initial ACT test. I find that, while non-disadvantaged boys who overperform on the initial ACT test are substantially more likely to take additional tests, responses to initial ACT performance are muted for girls and disadvantaged students.

Finally, I use a simple simulation exercise to show that closing the income gap in multiple test-taking behavior and the resulting score improvements, conditional on initial ACT score, race, and gender, would reduce the income gap in admissions-relevant test scores by roughly 7%. The remainder of the income gap in admissions-relevant test scores can be accounted for by disparities in initial ACT scores.

My results contribute to recent literature and ongoing policy debate over universal college admissions testing policies. In particular, I offer a potential explanation for why prior studies find relatively small impacts of universal college admissions testing policies on college enrollment. I show that many students, particularly non-disadvantaged students with higher academic achievement who are likely competitive college applicants, take multiple admissions tests. Thus, the benefit of gaining access to a single test may be diluted among students who do not take additional tests beyond the in-school ACT. Nonetheless, rank distortions are not large enough to outweigh the positive impacts of universal admissions testing policies on testing access, particularly among economically disadvantaged students. Roughly 40%

of disadvantaged students in my sample score above the threshold for admission to North Carolina public 4-year universities on their first ACT attempt. Thus, even if distortions from multiple test-taking decrease disadvantaged students' test score rank by 1 percentile on average and subsequently reduce their likelihood of admission to selective universities, universal ACT testing creates a pathway for disadvantaged students to access public universities.

Two policy avenues could be employed to reduce income gaps in multiple admissions test-taking. The first, proposed by [Vigdor and Clotfelter \(2003\)](#), is to penalize multiple testing attempts rather than superscoring and essentially considering students' highest scores across test attempts. Specific policies could include using only the in-school ACT score or averaging scores across all test attempts. While such policies would eliminate distortions arising from nonrandom selection into multiple test-taking, they could generate opposition from parents, students, and educators concerned about vulnerability to transitory shocks, such as having a "bad test day," and about behavioral responses to a more rigid, high-stakes testing environment. The second avenue, proposed by [Goodman et al. \(2020\)](#), is to further encourage multiple test-taking among disadvantaged students. Specific policies could include administering the in-school ACT earlier to allow more time for subsequent test attempts, increasing access to SAT fee waivers and testing centers, and providing information about the benefits of multiple test-taking. While such policy changes will not eliminate disparities in admissions test scores and college enrollment, addressing distortions induced by multiple test-taking has the potential to meaningfully benefit disadvantaged students in the college admissions process.

References

- ACT Research (2015). The Condition of College & Career Readiness 2015. Technical report, ACT Research.
- Altonji, J., Arcidiacono, P., and Maurel, A. (2016). The Analysis of Field Choice in College and Graduate School. In *Handbook of the Economics of Education*, volume 5, pages 305–396. Elsevier.
- Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.
- Avery, C., Shi, L., and Magouirk, P. (2025). Test-Optional College Admissions: ACT and SAT Scores, Applications, and Enrollment Changes. Working Paper w34260, National Bureau of Economic Research, Cambridge, MA.
- Bailey, M. and Dynarski, S. (2011). Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. Working paper, National Bureau of Economic Research, Cambridge, MA.
- Belley, P. and Lochner, L. (2007). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital*, 1(1):37–89.
- Bloem, M. D., Pan, W., and Smith, J. (2021). College entrance exam-taking strategies in Georgia. *Southern Economic Journal*, 88(2):587–627.
- Bond, T. N., Bulman, G., Li, X., and Smith, J. (2018). Updating Human Capital Decisions: Evidence from SAT Score Shocks and College Applications. *Journal of Labor Economics*, 36(3):807–839.
- Chetty, R., Deming, D. J., and Friedman, J. N. (2025). Diversifying society’s leaders? The determinants and causal effects of admission to highly selective private colleges. *The Quarterly Journal of Economics*.
- College Board (2015). SAT Score-Use Practices by Participating Institutions. Technical report, College Board.

- Cook, E. E. and Turner, S. (2019). Missed Exams and Lost Opportunities: Who Could Gain From Expanded College Admission Testing? *AERA Open*, 5(2):1–20.
- Foote, A., Schulkind, L., and Shapiro, T. M. (2015). Missed signals: The effect of ACT college-readiness measures on post-secondary decisions. *Economics of Education Review*, 46:39–51.
- Friedman, J. N., Sacerdote, B., Staiger, D. O., and Tine, M. (2025). Standardized Test Scores and Academic Performance at Ivy Plus Colleges. *AEA Papers and Proceedings*, 115:676–681.
- Frisancho, V., Krishna, K., Lychagin, S., and Yavas, C. (2016). Better luck next time: Learning through retaking. *Journal of Economic Behavior & Organization*, 125:120–135.
- Goodman, J., Gurantz, O., and Smith, J. (2020). Take Two! SAT Retaking and College Enrollment Gaps. *American Economic Journal: Economic Policy*, 12(2):115–158.
- Goodman, S. (2016). Learning from the Test: Raising Selective College Enrollment by Providing Information. *Review of Economics and Statistics*, 98(4):671–684.
- Harmston, M. T. and Crouse, J. (2016). Multiple testers: What do we know about them? ACT Research & Policy Brief, ACT Research.
- Hoxby, C. M. and Turner, S. (2015). What High-Achieving Low-Income Students Know About College. *American Economic Review*, 105(5):514–517.
- Hurwitz, M., Smith, J., Niu, S., and Howell, J. (2015). The Maine Question: How Is 4-Year College Enrollment Affected by Mandatory College Entrance Exams? *Educational Evaluation and Policy Analysis*, 37(1):138–159.
- Hyman, J. (2017). ACT for All: The Effect of Mandatory College Entrance Exams on Postsecondary Attainment and Choice. *Education Finance and Policy*, 12(3):281–311.
- Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., and Dunlop Velez, E. (2015). The condition of education 2015 (NCES 2015- 144). Technical report, Washington, DC: Department of Education, National Center for Education Statistics.

- Klasik, D. (2013). The ACT of Enrollment: The College Enrollment Effects of State-Required College Entrance Exam Testing. *Educational Researcher*, 42(3):151–160.
- Krishna, K., Lychagin, S., and Frisancho, V. (2018). Retaking in high stakes exams: Is less more? *International Economic Review*, 59(2):449–477.
- Lovell, D. and Mallinson, D. (2021). How Test-Optional College Admissions Expanded during the COVID-19 Pandemic. Brief, Urban Institute.
- McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., Barmer, A., Cataldi, E. F., and Mann, F. B. (2018). The Condition of Education 2018. NCES 2018-144. Technical report, National Center for Education Statistics.
- myFutureNC (2025). North Carolina ACT Performance Dashboard, 2024–25. Technical report, myFutureNC.
- Page, L. C. and Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, 51:4–22.
- Papay, J. P., Murnane, R. J., and Willett, J. B. (2016). The Impact of Test Score Labels on Human-Capital Investment Decisions. *Journal of Human Resources*, 51(2):357–388.
- Rosinger, K., Baker, D. J., Sturm, J., Yu, W., Park, J. J., Poon, O., Kim, B. H., and Breen, S. (2024). Exploring the Relationship between Test-Optional Admissions and Selectivity and Enrollment Outcomes during the Pandemic. EdWorkingPaper 24-982, Annenberg Institute for School Reform at Brown University.
- Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, 121(1-2):297–317.
- Sacerdote, B., Staiger, D. O., and Tine, M. (2025). How test optional policies in college admissions disproportionately harm high achieving applicants from disadvantaged backgrounds. Working Paper w33389, National Bureau of Economic Research, Cambridge, MA.

- Swiderski, T. (2025). Testing the way forward: The impact of statewide ACT or SAT testing on postsecondary outcomes. *Educational Evaluation and Policy Analysis*, 47(3):730–750.
- Vigdor, J. L. and Clotfelter, C. T. (2003). Retaking the SAT. *Journal of Human Resources*, 38(1):1–33.
- Westrick, P. A., Marini, J. P., Young, L., Ng, H., Shmueli, D., and Shaw, E. J. (2019). Validity of the SAT for Predicting First-Year Grades and Retention to the Second Year. Technical report, College Board.

Appendix

A Robustness of Multiple Test-Taking Regressions

Table A shows that the main results shown in Table 2 are robust to the inclusion of polynomials in ACT score and GPA in equations (1), (2), and (3). Second- and third-order polynomial terms are not statistically significant.

Table A2 shows that estimating equation (1) using a linear probability model instead of a probit model leads to estimates that are similar in magnitude. Translating probit coefficient estimates from Table 2 to average marginal effects suggests that disadvantaged students are 9.1 percentage points less likely to take multiple tests, while OLS coefficient estimates suggest a gap of 11.8 percentage points. Thus, the probit coefficient estimates shown in my main analysis are comparatively conservative.

Table A3 shows that the main rank change results shown in Table 2 are robust to converting all scores to the coarser 36-point ACT scale instead of the finer 1600-point SAT scale. This suggests that increases in percentile rank from taking the SAT in addition to the ACT are not simply driven by the SAT's finer 1600-point scale breaking ties between students with the same ACT scores. Main estimates from Table 2 suggests that disadvantaged students decrease in rank by 1.2 percentiles, on average, when ranking students on their admissions-relevant test scores instead of their initial ACT scores, while estimates using the 36-point ACT scale in Table A3 suggest a decrease of 1.3 percentiles.

Table A1: Multiple Test-Taking and Student Characteristics, Polynomial Specification

	(1)	(2)	(3)
	Took Multiple Tests	Score Change	Rank Change
ACT Score	-0.0000948 (0.000685)	-2.240*** (0.0647)	0.00347 (0.00391)
10th Grade GPA	0.767*** (0.0534)	57.48*** (5.369)	1.034*** (0.213)
Female	0.135*** (0.00531)	-19.70*** (0.353)	-0.775*** (0.0297)
Black	0.431*** (0.0118)	-6.755*** (0.593)	1.312*** (0.0672)
Economically disadvantaged (EDS)	-0.427*** (0.00922)	-5.717*** (0.550)	-1.204*** (0.0388)
Black \times EDS	0.265*** (0.0135)	-2.323** (0.856)	0.0597 (0.0786)
Hispanic	-0.214*** (0.0137)	-4.741*** (0.787)	-1.031*** (0.0674)
Hispanic \times EDS	0.289*** (0.0172)	-2.713* (1.096)	0.415*** (0.0739)
Asian	0.260*** (0.0215)	20.33*** (1.682)	1.460*** (0.123)
Asian \times EDS	0.151*** (0.0364)	-17.41*** (2.352)	-0.810*** (0.204)
Model	Probit	OLS	OLS
N	380465	177273	381599
R^2 /Pseudo- R^2	0.288	0.256	0.137

Coefficients from probit and OLS regressions corresponding to equations (1), (2), and (3)

All models include school and cohort fixed effects

Models include cubic polynomials in ACT score and GPA

Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2: Multiple Test-Taking and Student Characteristics, Linear Probability Model

	(1) Took Multiple Tests
ACT Score	0.000469*** (0.00000790)
10th Grade GPA	0.179*** (0.00164)
Female	0.0418*** (0.00151)
Economically disadvantaged (EDS)	-0.118*** (0.00258)
Black	0.123*** (0.00352)
Black \times EDS	0.0556*** (0.00386)
Hispanic	-0.0647*** (0.00395)
Hispanic \times EDS	0.0775*** (0.00466)
Asian	0.0500*** (0.00499)
Asian \times EDS	0.0569*** (0.0104)
Model	OLS
N	381599
R^2	0.332

Coefficients from OLS regression corresponding to equation (1)

Model includes school and cohort fixed effects

Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3: Rank Change and Student Characteristics, 36-Point ACT Scale

	(1) Rank Change
ACT Score	-0.0172*** (0.000240)
10th Grade GPA	3.040*** (0.0404)
Female	-0.765*** (0.0315)
Economically disadvantaged (EDS)	-1.328*** (0.0412)
Black	1.530*** (0.0715)
Black \times EDS	0.106 (0.0833)
Hispanic	-1.088*** (0.0727)
Hispanic \times EDS	0.496*** (0.0807)
Asian	1.773*** (0.125)
Asian \times EDS	-0.947*** (0.217)
Model	OLS
N	381599
R^2	0.138

Coefficients from OLS regression corresponding to equation (3)

Model includes school and cohort fixed effects

Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

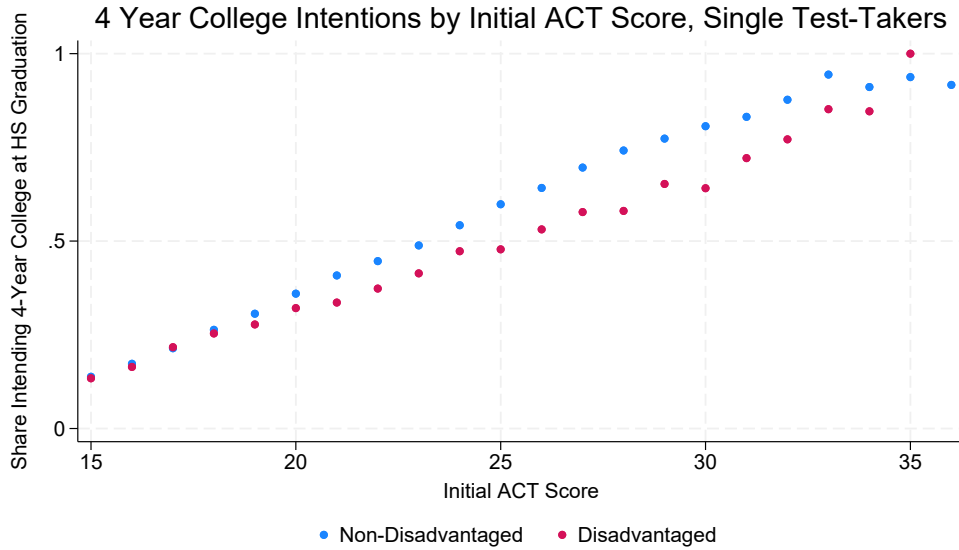
B Multiple Test-Taking and College Enrollment

While the NCERDC data does not include any details on the specific colleges students plan to attend, I observe a measure of intended 4-year college enrollment from a graduation survey administered during the spring of 12th grade. Descriptive evidence indicates that 4-year college enrollment is substantially higher among students who take multiple college admissions tests: 76% of students who take multiple tests intend to enroll in a 4-year college compared to 23% of students who take only the in-school ACT.

Figures A1 and A2 demonstrate that the income gap in 4-year college enrollment is slightly smaller among multiple test-takers when comparing students with the same initial ACT scores. Higher college enrollment rates and smaller income gaps in college enrollment among multiple test-takers could be due to positive selection on motivation and college intent or to gains in test scores and test score rank from multiple test-taking.

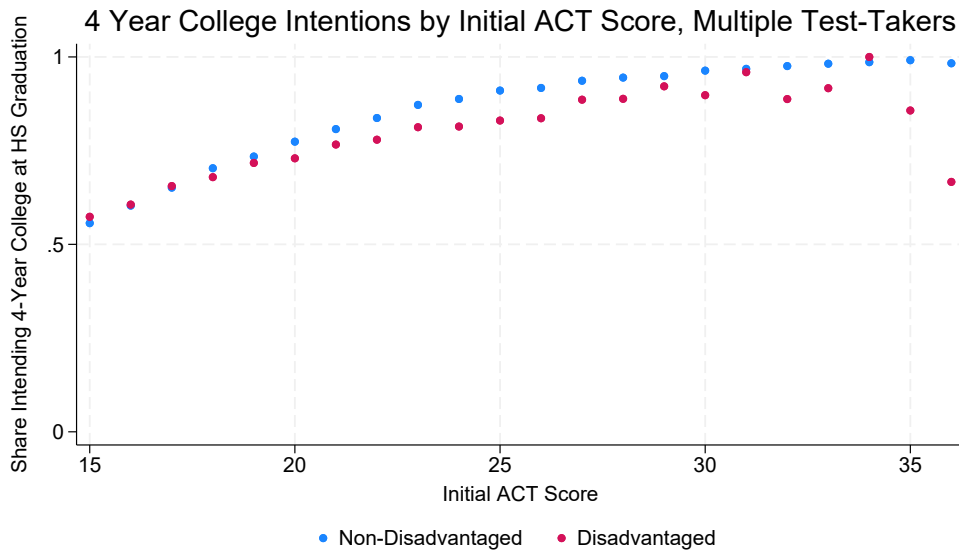
Figure A3 further plots income gaps in intended 4-year college enrollment by initial ACT score among multiple test-takers who do not improve upon their initial score. While the average student who takes multiple tests improves on their initial score by 56 points on the 1600-point SAT scale, 37% of multiple test-takers receive an equivalent or lower score from subsequent test attempts. These students' maximum test scores are equivalent to their initial ACT scores. Among these students who select into multiple test-taking but do not experience test score or rank gains, the income gap in 4-year college enrollment is similar in size to the income gap among the full sample of multiple test-takers. This pattern suggests that the smaller income gap in 4-year college enrollment observed among multiple test-takers, relative to single test-takers, is primarily driven by selection into multiple test-taking. Nonetheless, test score and rank gains from multiple test-taking likely contribute to income gaps in college enrollment on the margin.

Figure A1: College Enrollment by Initial ACT Score and EDS, Single Test-Takers



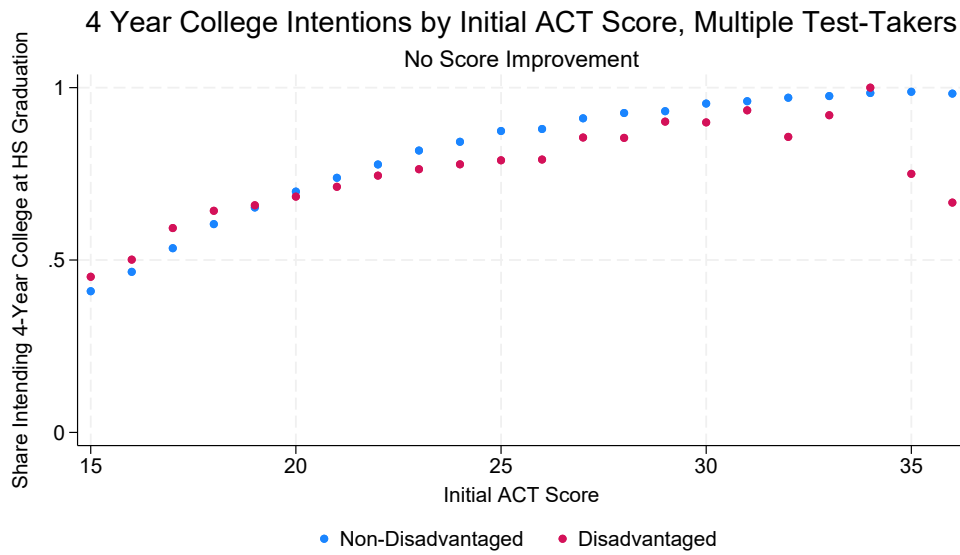
Notes: Means of intended 4-year college enrollment by initial ACT score, calculated from the raw data. Sample restricted to students with an initial ACT score of 15 or higher due to high variability in college enrollment at lower scores.

Figure A2: College Enrollment by Initial ACT Score and EDS, Multiple Test-Takers



Notes: Means of intended 4-year college enrollment by initial ACT score, calculated from the raw data. Sample restricted to students with an initial ACT score of 15 or higher due to high variability in college enrollment at lower scores.

Figure A3: College Enrollment by Initial ACT Score and EDS, Multiple Test-Takers with No Score Improvement

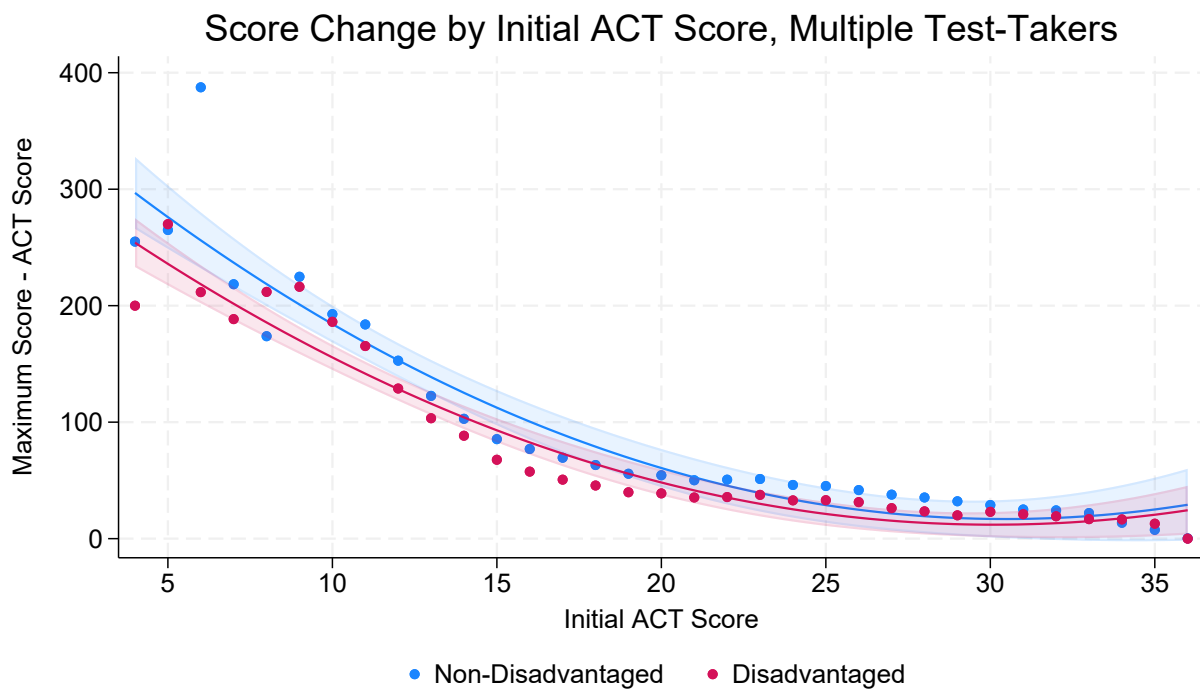


Notes: Means of intended 4-year college enrollment by initial ACT score, calculated from the raw data. Sample restricted to students with an initial ACT score of 15 or higher due to high variability in college enrollment at lower scores.

C Score Improvements by Initial ACT Score

Figure A4 demonstrates that average disparities in score changes by disadvantaged status are small within initial ACT scores. This result suggests that large increases in rank for disadvantaged multiple test-takers relative to non-disadvantaged multiple test takers arise from who chooses to retest, rather than from systematically larger gains to multiple test-taking among disadvantaged students.

Figure A4: Mean Score Improvement by Initial ACT Score and EDS



D Robustness of Multiple Test-Taking and Belief Updating

Table A4 shows that estimating equation (10) using a linear probability model instead of a probit model leads to estimates that are similar in magnitude. Translating main probit coefficient estimates from Table 3 to average marginal effects suggests that a one standard deviation increase in ACT score residual increases multiple test-taking by roughly 2 percentage points among non-disadvantaged males. OLS coefficient estimates, divided by the standard deviation of the residual of 2.26, suggest a slightly smaller increase of 0.98 percentage points.

Table A5 shows that estimating equation (10) among the full sample, rather than restricting the sample to students with predicted ACT scores in the middle 50% of the predicted score distribution, leads to results that are qualitatively similar but smaller in magnitude. This is likely because most students with low academic achievement do not take multiple admissions tests and most students with very high academic achievement take multiple admissions tests regardless of their initial ACT performance, in order to be competitive for selective college admissions.

Table A6 shows that effects are robust to including the full set of ACT score predictors in the regression, rather than summarizing students' expected ACT performance using the predicted ACT score measure. Translating coefficient estimates to average marginal effects suggests that a one standard deviation increase in ACT score residual increases multiple test-taking by roughly 1 percentage point among non-disadvantaged males.

Table A4: Multiple Test-Taking and ACT Score Residual, Linear Probability Model

	(1)	(2)
	Female	Male
Predicted ACT Score	0.0506*** (0.00107)	0.0480*** (0.00106)
ACT Score Residual	0.0154*** (0.00108)	0.0222*** (0.000901)
Economically disadvantaged	-0.0725*** (0.00383)	-0.0672*** (0.00374)
Economically disadvantaged \times Residual	-0.00480** (0.00163)	-0.00681*** (0.00134)
Model	OLS	OLS
N	80066	77206
R^2	0.158	0.152

Coefficients from OLS regressions corresponding to equation (10)

All models include school and cohort fixed effects

Sample includes students in the 25th-75th percentiles of predicted ACT scores

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A5: Multiple Test-Taking and ACT Score Residual, Full Sample

	(1)	(2)
	Female	Male
Predicted ACT Score	0.131*** (0.00149)	0.132*** (0.00138)
ACT Score Residual	0.0227*** (0.00242)	0.0389*** (0.00201)
Economically disadvantaged	-0.213*** (0.00854)	-0.200*** (0.00927)
Economically disadvantaged \times Residual	0.000254 (0.00352)	0.00257 (0.00321)
Model	Probit	Probit
N	158121	154934
Pseudo- R^2	0.187	0.200

Coefficients from probit regressions corresponding to equation (10)

All models include school and cohort fixed effects

Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: Multiple Test-Taking and ACT Score Residual, Full Set of ACT Score Predictors

	(1)	(2)
	Male	Female
ACT Score Residual	0.0253*** (0.00243)	0.0376*** (0.00213)
Economically disadvantaged	-0.116** (0.0378)	-0.0922* (0.0404)
Economically disadvantaged \times Residual	0.0103** (0.00367)	0.00878* (0.00350)
Model	Probit	Probit
N	158121	154934
Pseudo- R^2	0.257	0.280

Coefficients from probit regressions corresponding to equation (10)

All models include school and cohort fixed effects

Models include controls from equation (9)

Sample includes students in the 25th-75th percentiles of predicted ACT scores

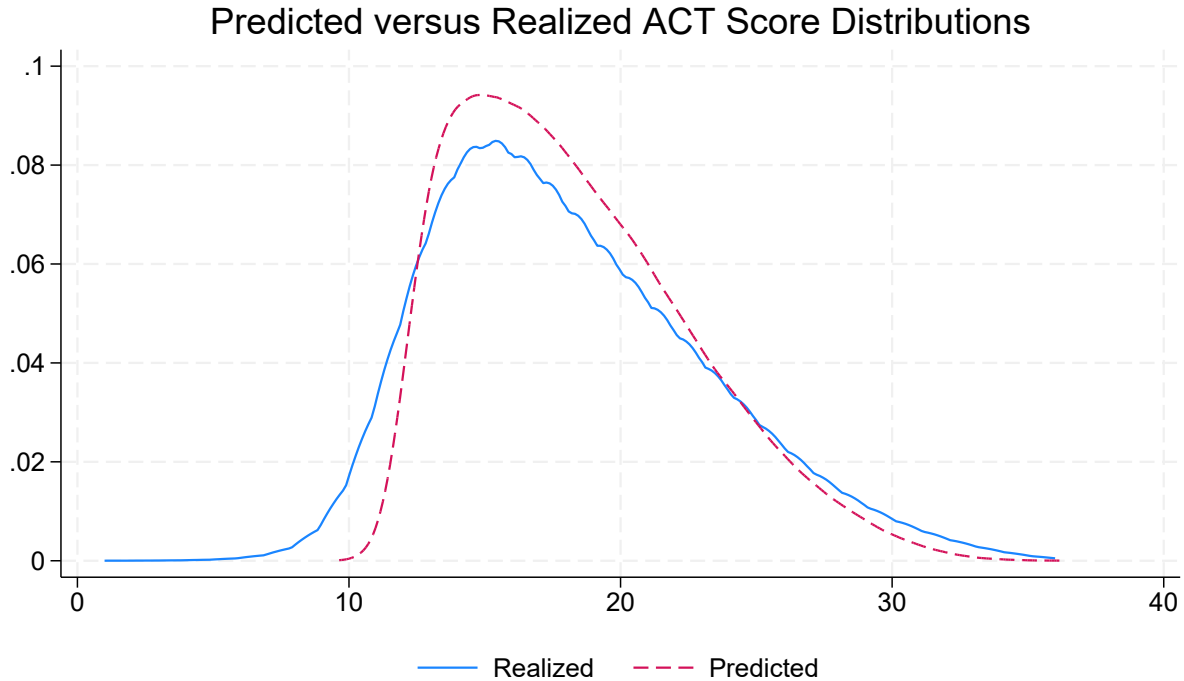
Standard errors in parentheses, clustered at the school-cohort level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

E ACT Score Prediction Model Fit

Figure A5 compares the predicted ACT score distribution with the true ACT score distribution. The prediction model fits the shape of the ACT score distribution well with an R^2 of 0.782, despite slightly underpredicting mass at the upper and lower tails.

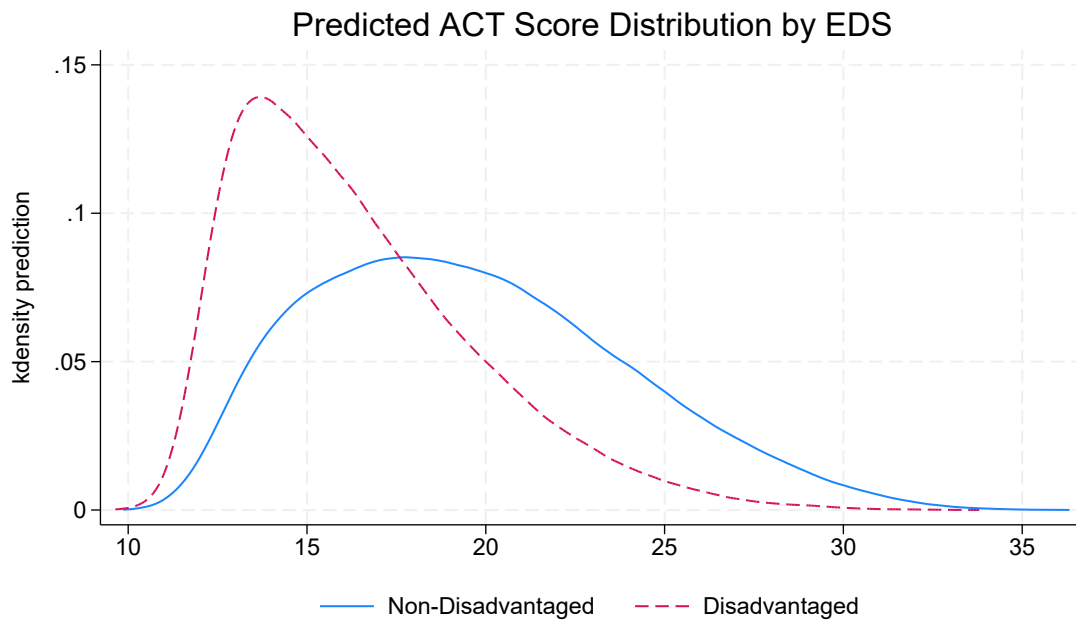
Figure A5: Prediction Model Fit



Notes: Predicted ACT scores \hat{ACT}_{ist} , on the 36-point ACT scale, from equation (9)

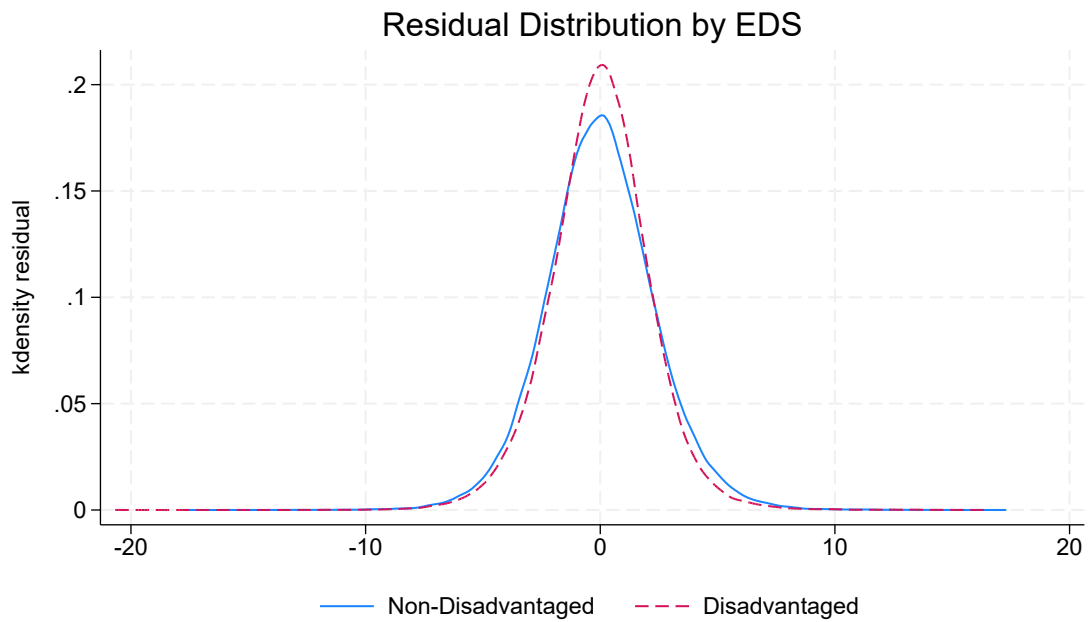
A positive value of the residual indicates that a student overperformed on the ACT test relative to his or her past performance, and vice versa. The standard deviation of the residual is 2.26. The standard deviation of both the predicted ACT score and the ACT score residual is slightly smaller among disadvantaged students, as demonstrated in Figures A6 and A7, respectively.

Figure A6: Distribution of Predicted ACT Score \hat{ACT}_{ist}



Notes: Predicted ACT scores \hat{ACT}_{ist} , on the 36-point ACT scale, from equation (9)

Figure A7: Distribution of Residual $\hat{\epsilon}_{ist}$ By EDS



Notes: Residuals $\hat{\epsilon}_{ist}$ from equation (9)