

ACT FOR ALL: THE EFFECT OF MANDATORY COLLEGE ENTRANCE EXAMS ON POSTSECONDARY ATTAINMENT AND CHOICE

Joshua Hyman

Department of Public Policy
University of Connecticut
West Hartford, CT 06117
joshua.hyman@uconn.edu

Abstract

This paper examines the effects of requiring and paying for all public high school students to take a college entrance exam, a policy adopted by eleven states since 2001. I show that prior to the policy, for every ten poor students who score college-ready on the ACT or SAT, there are an additional five poor students who would score college-ready but who take neither exam. I use a difference-in-differences strategy to estimate the effects of the policy on postsecondary attainment and find small increases in enrollment at four-year institutions. The effects are concentrated among students less likely to take a college entrance exam in the absence of the policy and students in the poorest high schools. The students induced by the policy to enroll persist through college at approximately the same rate as their inframarginal peers. I calculate that the policy is more cost-effective than traditional student aid at boosting postsecondary attainment.

doi:10.1162/EDFP_a_00206

© 2017 Association for Education Finance and Policy

1. INTRODUCTION

Inequality in educational attainment has widened substantially during recent decades. Not only do minority and low-income students enroll in postsecondary education in lower proportions than their majority and higher-income counterparts, but conditional on enrolling, these students are less likely to persist through college and complete a degree (Bailey and Dynarski 2011). Although certainly not every low-income and minority student would benefit from postsecondary education, recent research suggests that a nontrivial number of high-achieving, disadvantaged students either do not attend college or attend a less selective school than they could (Pallais and Turner 2006; Bowen, Chingos, and McPherson 2009; Hoxby and Avery 2013; Dillon and Smith 2017). Policies that induce low-income students to attend and persist at appropriately selective institutions could have substantial implications for reducing educational inequality.

Many policies and interventions aim to increase the educational attainment of disadvantaged students. Policies such as Head Start, class size reduction, and school finance reform, which aim to increase the human capital of students, as well as policies such as student aid that reduce the cost of college, have all been shown to successfully increase postsecondary attainment (Deming 2009; Deming and Dynarski 2010; Dynarski, Hyman, and Schanzenbach 2013; Hyman, forthcoming). These policies are all quite expensive, however, costing tens of thousands of dollars to induce one additional student to enroll in college (Dynarski, Hyman, and Schanzenbach 2013). Recently, interventions aimed at reducing informational and administrative barriers to college enrollment have found large effects at a fraction of the cost of the more traditional tools mentioned above (Bettinger et al. 2012; Hoxby and Turner 2012; Carrell and Sacerdote forthcoming). It remains to be seen whether these low-cost policies can be implemented effectively at scale.

In this paper, I examine the impacts of an inexpensive policy aimed at boosting postsecondary attainment that is currently operating at scale. Eleven states require and pay for college entrance exams (i.e., the ACT or SAT) for all public school eleventh graders. Given that it costs less than \$50 per student for states to implement this policy, very small effects on college-going would suffice for the policy to be as cost effective as traditional student aid. In this paper, I examine the effect of mandatory college entrance exams on postsecondary enrollment, persistence, and choice. I use an original student-level dataset containing six complete cohorts of eleventh-grade public high school students in Michigan, a state that implemented a mandatory ACT policy in 2007. The data include demographics, eighth- and eleventh-grade statewide assessment scores, information on postsecondary enrollment, and ACT and SAT scores for all test-takers during the sample period.

To begin my analysis, I use the post-policy ACT score distribution to deduce what fraction of pre-policy non-takers would score at a college-ready level if they took the exam.¹ I show that for every ten poor students taking a college entrance exam and scoring college-ready, there are an additional five poor students who do not take the test but who would score college-ready if they did. This represents a contribution to

1. The basic intuition for how I calculate the pre-policy number of students who did not take the exam, but would have scored college-ready, is by subtracting the number of test-takers who score college-ready in the pre-period from the number who do so in the post-period.

the emerging literature on “undermatch.” Hoxby and Avery (2013) focus on the supply of disadvantaged students who take a college entrance exam and score in the top 10 percent of takers but do not apply to selective colleges. I use a lower threshold of “high-achieving,” and look back further in the college application process, finding a large supply of disadvantaged students who would score well enough to enroll in a selective four-year college but who are dropping out of the application process prior to even taking a college-entrance exam.

To examine the effects of the mandatory ACT policy on postsecondary outcomes, I use a difference-in-differences (DID) style approach that compares changes in college-going from before to after the implementation of the policy for students in schools without a test center in the pre-policy period relative to students in schools that had a test center. In doing so, I exploit the fact that schools without a test center pre-policy had lower test-taking rates and thus experience a larger treatment dosage. I use propensity score matching to restrict my analysis to a sample of test center and non-test center schools that have similar observed characteristics.

I estimate a 0.6 percentage point (2 percent) effect of the policy on the probability that a student enrolls in a four-year college. This overall effect masks important heterogeneity, with larger effects (1.3 points, 5 percent) for students with a low-to-mid-level probability of taking the ACT in the absence of the policy. Effects are also larger among males (0.9 points, 3 percent), poor students (1.0 points, 6 percent), and students at schools with a high poverty share (1.3 points, 6 percent). Two recent studies estimate the effects of the mandatory ACT policy using aggregate state-level data, and thus cannot estimate heterogeneity by student or school characteristics (Klasik 2013; Goodman 2016). By using microdata, I am able to show that this policy is in fact effective at reducing inequality, with effects on college enrollment concentrated among economically disadvantaged students and poor schools.

Finally, I find suggestive evidence that the marginal student induced into college by the policy persists through college at the same rate as the inframarginal student. Because my data follow students over time, my study can estimate persistence through college as a result of the policy. Given the extent of inequality in postsecondary persistence (Bailey and Dynarski 2011), this is a necessary parameter for understanding the policy’s full welfare effects.

The most similar study to my own is that by Hurwitz et al. (2015), which uses College Board microdata and a DID approach to estimate the four-year college enrollment effects of Maine’s mandatory SAT policy.² The present paper makes two primary contributions beyond Hurwitz et al. The first is external validity: Maine is a small and unique state, whereas Michigan is a large and more representative state. Further, most state-mandated college entrance exam policies require the ACT and are offered during normal school hours. The Maine policy requires the SAT and is offered only on Saturday. To the extent that these policy features alter the policy’s effects, the Michigan case may be more generalizable. The second contribution is that because of data limitations, Hurwitz et al. are unable to estimate effects on two-year college enrollment. I show that the policy’s effect on four-year college enrollment is not primarily due to displacing two-year enrollments.

2. I compare the results of that study with my own in section 6.

The DID estimator used in this paper yields an effect that is arguably causal but is a lower bound of the true policy impact because some portion of the effect is likely experienced equally by students at both test center and non-test center schools, and is thus not captured by this methodology. Using this lower bound, however, I calculate that the mandatory college entrance exam policy is more cost-effective than traditional student aid at boosting postsecondary attainment.

The remainder of this paper is structured as follows: Section 2 discusses the mandatory college entrance exam policy. Section 3 describes the data. Section 4 examines the population of college-ready students not taking a college entrance exam pre-policy. Section 5 examines the policy's effects on postsecondary outcomes. Section 6 discusses the interpretation of my DID estimates and possible supply-side capacity constraints. Finally, section 7 concludes with a comparison of the costs and benefits of mandatory college entrance exams to other education policies.

2. COSTS, INFORMATION, AND MANDATORY COLLEGE ENTRANCE EXAMS

The ACT and SAT are college admission exams required for admission to nearly all four-year institutions across the country.³ Historically, these exams have been taken exclusively by students considering applying to a four-year institution. Since 2001, however, eleven states have implemented free and mandatory college entrance exams for all high school juniors, and several more are planning to implement the reform in the near future.⁴ These states tend to cite increasing college access as the motivation for the policy. Most of the mandatory ACT-adopting states are centrally located within the United States in the Central and Mountain census divisions. After Illinois, Michigan is the most populous state to have adopted the policy.

The state-mandated ACT and SAT are the official exams used for college admission purposes. Traditionally, the ACT and SAT are offered on Saturday mornings, cost students between \$30 and \$50, and require students to travel to the nearest test center. Fee waivers are available for low-income students but take-up is low, perhaps because it requires paperwork on the part of the student and coordination with high school counselors. State-mandated exams are typically given during the school day, at no financial cost to the student, and at the student's high school. As with the standard ACT and SAT, students can select colleges to which they send their scores. Students are mailed an official score report several weeks after they take the exam. Mandatory college entrance exams provide a substantial change to the structure of the four-year college application process that reduces the monetary, psychic, and time cost of applying to college.⁵ While spending \$30 to \$50 and five hours on a Saturday represents a small share of the overall

3. Exceptions are primarily for-profit institutions, specialty or religious institutions, and institutions that admit all or nearly all applicants. All four-year public universities in Michigan require the ACT or SAT for admission.
4. Appendix table A.1 (which can be found on the *Education Finance and Policy* Web site at www.mitpressjournals.org/doi/suppl/10.1162/EDFP_a_00206) lists the states that have adopted this policy, which exam they use (nearly all use the ACT rather than the SAT), and the year that the first eleventh grade cohort was exposed to the policy. In order of adoption, the states are: Colorado, Illinois, Maine, Michigan, Kentucky, Tennessee, Delaware, North Carolina, Louisiana, Wyoming, and Alabama.
5. Recent research has shown that small changes to the structure of choice-making, such as changes in the default choice, can have large behavioral effects in various policy domains like retirement savings plans (Madrian and Shea 2001; Beshears et al. 2009). Similarly, a small change to the structure of the college entrance exam score report sending process was shown to have large effects on the number of score reports students sent (Pallais 2015).

cost of applying to and attending college, these monetary and time costs can represent a real hurdle to low-income students, particularly if taking the test requires seeking time off from employment. Further, approximately half of public school students do not attend a high school with a test center in the school, so they would have to find and travel to the nearest test center.⁶ Offering the exam for free during school all but eliminates these costs to the student.

Mandatory college entrance exams could also alleviate information constraints in the college application process. Students taking the ACT or SAT may learn about college accessibility because after the test they may receive mailings from postsecondary institutions. Test-takers may also learn about their college-going ability. The score on these tests provides students with a signal of their likelihood of being admitted to, and succeeding at, a four-year college or university.

Finally, mandatory college entrance exams may increase information about the college application process by altering school-level behavior. In Michigan, most schools have at least some resources available to help students prepare for the tests, and some schools with greater resources offer entire classes devoted to preparing for the exams.⁷ More broadly, this policy has the potential to increase the college-going culture at a school, which has been shown to be an important instrument in increasing the postsecondary attainment of disadvantaged students (Jackson 2010).

3. DATA

This paper uses an original dataset containing all students attending Michigan public high schools in six recent eleventh grade cohorts (2003–04 through 2008–09). The data contain time-invariant demographics such as sex, race, and date of birth, as well as time-varying characteristics such as free and reduced-price lunch status, limited-English-proficiency (LEP) status, special education (SPED) status, and student's home address. The data also contain eighth and eleventh grade state assessment results. For the cohorts of students exposed to the mandatory ACT exam, the eleventh-grade assessment results include ACT scores. Student-level postsecondary enrollment information is obtained by matching students to the National Student Clearinghouse (NSC).⁸ School- and district-year level characteristics from the Common Core of Data are merged to the dataset based on where and when students are enrolled in high school.

I acquired and merged in several other key pieces of information. Using student name, date of birth, sex, race, and eleventh grade home zip code, I matched the Michigan data to microdata from ACT, Inc., and The College Board on every ACT- and SAT-taker in Michigan over the sample period. This allows me to observe ACT-takers pre-policy, as well as students who took the SAT instead of the ACT pre-policy. I also

6. Bulman (2015) finds that the opening of an SAT test center in a high school has large effects on SAT-taking, and on educational attainment. That paper also examines the effects in three school districts (Stockton, CA; Palm Beach, FL; and Irving, TX) of offering a free SAT. He finds four-year enrollment effects of the policies on the order of 15 percent. Although these effects are larger than those I estimate, a single district in the state offering the SAT for free is quite a different policy than a statewide implementation of a mandatory exam.

7. From author's discussions with guidance counselors and state departments of education.

8. The NSC is a nonprofit organization that houses postsecondary enrollment information on over 90 percent of undergraduate enrollment nationwide. See Dynarski, Hemelt, and Hyman (2015) for a detailed discussion of the NSC matching process and coverage rates.

Table 1. Sample Means of Michigan Eleventh Grade Student Cohorts

	All Cohorts (2004–09) (1)	Pre-ACT Cohorts (2004–06) (2)	Post-ACT Cohorts (2007–09) (3)	Difference: (3) – (2) (4)	p-Value: (4) = 0 (5)
Demographics					
Female	0.498	0.498	0.498	–0.001	0.436
White	0.764	0.778	0.751	–0.027	0.000
Black	0.167	0.155	0.179	0.024	0.000
Hispanic	0.033	0.031	0.035	0.004	0.000
Other race	0.035	0.036	0.035	0.000	0.258
Free or reduced lunch	0.283	0.241	0.322	0.080	0.000
Special education	0.123	0.124	0.122	–0.002	0.041
Limited English	0.021	0.020	0.023	0.002	0.000
Local unemployment	8.25	7.34	9.13	1.79	0.000
Driving miles to nearest ACT test center	3.60	4.72	2.52	–2.20	0.000
Educational attainment					
Reaches twelfth grade	0.908	0.904	0.912	0.009	0.000
Graduates high school	0.844	0.844	0.844	0.000	0.923
Enrolls in any college	0.580	0.570	0.589	0.020	0.000
Enrolls in four-yr college	0.314	0.309	0.319	0.010	0.000
ACT score	19.6	20.7	18.9	–1.9	0.000
ACT-taking rate					
All students	0.739	0.558	0.912	0.354	0.000
Males	0.706	0.507	0.898	0.392	0.000
Females	0.771	0.611	0.926	0.315	0.000
Blacks	0.647	0.456	0.806	0.350	0.000
Whites	0.761	0.583	0.939	0.355	0.000
Free or reduced lunch	0.644	0.350	0.852	0.503	0.000
Non-free lunch	0.778	0.625	0.940	0.316	0.000
<Median grade eight score	0.662	0.401	0.902	0.501	0.000
>Median grade eight score	0.868	0.766	0.961	0.195	0.000
Missing grade eight score	0.101	0.123	0.079	–0.044	0.000
Took SAT	0.048	0.064	0.033	–0.030	0.000
Took SAT & ACT	0.045	0.058	0.033	–0.025	0.000
SAT score	25.0	24.7	25.8	1.1	0.000
Students per cohort	122,243	119,917	124,570		
Total students	733,460	359,751	373,709		

Notes: The sample is all first-time eleventh graders in Michigan public high schools during 2003–04 through 2008–09 conditional on reaching their eleventh-grade spring semester. Free lunch, special education, and limited English proficiency status are all as of eleventh grade. Driving miles to nearest ACT test center are measured from a student's home address during eleventh grade to the nearest ACT test center open during that year. First score is used for students taking the ACT multiple times. SAT score is scaled to ACT metric. College enrollment is measured as of 16 months (1 October) following scheduled on-time high school graduation. Eighth grade score is the average of scores on the eighth-grade math and writing exams, standardized at the subject-cohort level.

acquired from ACT, Inc., a list of all ACT test centers in Michigan over the sample period, including their addresses and open and close dates. For a robustness check, I geocoded student home addresses during eleventh grade, and the addresses of these test centers, to calculate the driving distance from the student's home to the nearest center.

Table 1 shows sample means before and after implementation of the mandatory ACT. I condition my sample on reaching the spring semester of eleventh grade, which

is the semester when the eleventh-grade state assessment is given. Michigan was hit hard by the economic recession during the sample period: The percentage of eleventh graders eligible for free lunch rose from 24 percent to 32 percent, and the percentage that are black increased from 15.5 percent to 18 percent. The local city- (if available) or county-level unemployment rate obtained from the Bureau of Labor Statistics rose from 7.3 percent to 9.1 percent. Educational attainment was fairly stable over the period, with high school graduation at 84.4 percent and college enrollment increasing slightly from 57 percent to 59 percent.⁹

Prior to the mandatory ACT policy, 56 percent of students took the ACT. The percentage increased to 91 percent after the policy. ACT-taking rates tend to increase more for those groups of students who have lower rates prior to the policy. This is particularly pronounced among students eligible for free or reduced-price lunch, whose rate of ACT-taking more than doubles, from 35 percent to 85 percent.

High school dropout is the primary source of noncompliance during the post-policy period. Of the 91 percent of students in my sample who reach twelfth grade, the test-taking rate is 95 percent. Of those who graduate high school, the test-taking rate is 97.8 percent. The remaining noncompliance is mostly due to students taking the special education version of the eleventh grade test that does not include the ACT. Of the 80 percent of the sample who are high school graduates and who do not take the special education version of the test, the fraction with a valid ACT score is 98.9 percent.

4. SELECTION INTO COLLEGE ENTRANCE EXAM-TAKING

In this section, I use the implementation of the policy as a natural experiment that allows me to measure the pre-policy supply of students who did not take a college entrance exam but would have scored well had they taken one. The intuition behind the framework that I develop is that I treat the post-policy ACT score distribution as the distribution that would be observed pre-policy had all students taken the exam. Under a set of assumptions discussed below, this allows me to recover the latent distribution of test scores for students who did not take the test.

The Supply of High-achieving Non-takers

I begin my analysis by predicting the ACT score distribution that would be observed among non-takers during the pre-policy period if they were to take the ACT. I do this by subtracting the number of test-takers scoring at each ACT score during the pre-policy period from the number scoring at each score in the post period, when nearly all students take the test.¹⁰ This simple strategy will recover the latent score distribution of all pre-policy non-takers under the assumptions that (1) the average size of the cohorts is the same pre- and post-policy, (2) the composition of public school students and

9. I define a student as enrolling in college if he or she enrolls before 1 October of the second fall following on-time high school graduation. This definition ensures that the measure is consistent across cohorts as I do not observe more than two years of enrollment for the most recent cohort. This variable can be thought of as a liberal measure of on-time college enrollment that captures students graduating high school on time and taking a gap year before enrolling, or students who take an extra year to graduate high school and then enroll the following fall.

10. For students taking the ACT multiple times, I use their first score. For pre-policy students who took the SAT but not the ACT, I include their SAT score scaled to the ACT metric. For students taking both tests, I use their first ACT score.

other factors in Michigan affecting ACT scores is stable over the sample period, and (3) all students take the ACT in the post period. As we have already seen, none of these assumptions is strictly true, so I adjust my procedure in a number of ways.

To ensure that the changing cohort size and composition of Michigan students is not leading to differences in the score distributions, I reweight the post-policy cohorts of students following DiNardo, Fortin, and Lemieux (1996) to resemble the pre-policy students according to their observed characteristics. Specifically, I estimate using ordinary least squares (OLS):

$$PRE_{isd} = \alpha + \beta_1 X_{isdt} + \beta_2 S_{st} + \beta_3 D_{dt} + \varepsilon_{isdt}, \quad (1)$$

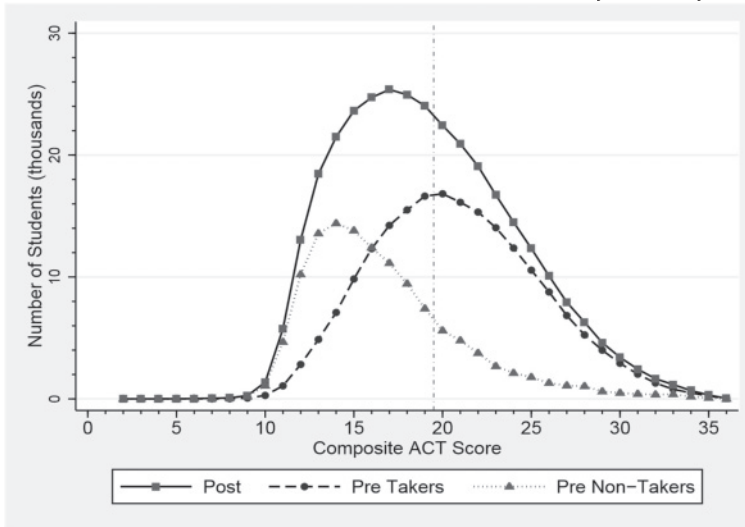
where PRE_{isd} is an indicator for student i in school s in district d being in the pre-policy period. X is a vector of individual-level covariates, S is a vector of school-year level covariates, and D is a vector of district-year level covariates.¹¹ I predict \widehat{PRE}_{isd} , which is the propensity score of being in the pre-policy period. The DiNardo, Fortin, and Lemieux (henceforth, DFL; 1996) weight equals $\frac{PRE_{is}}{(1 - \widehat{PRE}_{is})}$, which I then censor at its 1st and 99th percentile.¹² When adjusting the distribution, each pre-policy score receives a weight of 1, and each post-policy score receives its censored DFL weight. To adjust for increasing cohort size, I normalize the DFL weights in the post-policy period to have a mean equal to 0.963, which is the proportional size of the three combined pre-policy cohorts relative to the three combined post-policy cohorts. To compute the distribution of latent scores, I sum the weights in the post period at each ACT score, and subtract the sum of the weights at each score in the pre-period.

Panel A of figure 1 shows this exercise graphically: the dashed line plots the frequency distribution of scores pre-policy, which is skewed slightly to higher achievers. The solid line plots the reweighted post-policy score distribution, which is larger because there are many more test-takers, and is substantially skewed to low achievers, reflecting the lower average scores of students induced into test-taking. Assume that after DFL-reweighting the only difference between the pre- and post-policy cohorts is that nearly everyone takes the ACT in the post period. Then the difference in the number of students scoring at each ACT score bin should reflect the distribution of unobserved latent scores of the students who did not take the exam before it was mandatory.¹³

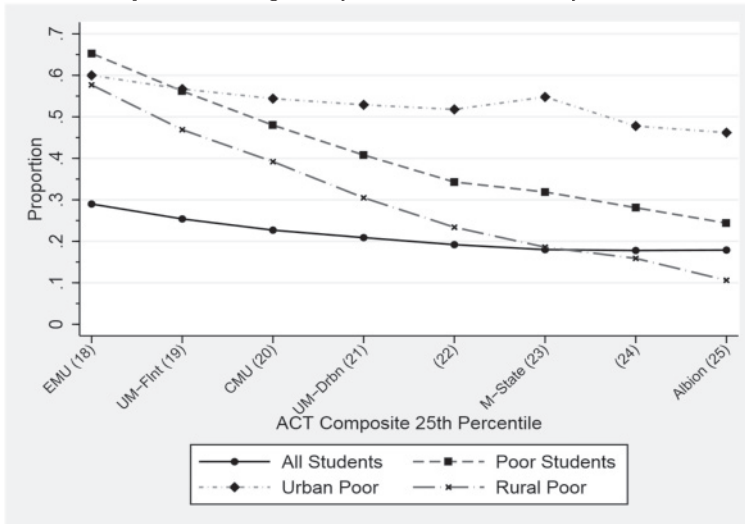
While the latent scores of pre-policy non-takers (figure 1 panel A, dotted line) are generally lower than the scores of those taking the test (dashed line), there is a long right tail of students who do not take a college entrance exam pre-policy, but would score college-ready if they did.¹⁴ As a threshold of college-readiness, I use a score of

11. X includes LEP, SPED, free lunch, race dummies, and sex. S includes fraction free lunch, fraction black, number of eleventh graders, and pupil–teacher ratio. D includes district-level versions of the variables in S plus student–counselor ratio, dummies for urban/rural status, and the local unemployment rate. All interactions of student-level covariates with each other and with the school- and district-year level covariates are included. The R^2 from the regression is 0.149.
12. I censor the weights because extremely low or high values of the DFL weight can be problematic (DiNardo 2002). In practice, the results are not sensitive to censoring the weights.
13. In practice, there is little difference between the results with and without the DFL-reweighting. Because the post-policy sample has a higher fraction minority and free-lunch eligible students, the DFL-reweighting places slightly higher weight on white and non-free lunch eligible students, slightly shifting the post-policy distribution upward.
14. A Kolmogorov–Smirnov nonparametric test of equality of the distribution of the observed scores among takers and the latent scores among non-takers is rejected with a p -value of 0.000. Appendix table A.2 (in the online

Panel A: ACT Score Distributions Pre- and Post-Mandatory ACT Policy



Panel B: Proportion of College-Ready Non-Takers to Takers, By ACT Score Threshold



Notes: Panel A shows the distribution of ACT scores post-policy, pre-policy, and the difference, which is the latent score distribution among non-takers in the pre-period. The pre-reform proportion of college-ready students not taking the ACT to those who take the ACT is 0.227. Panel B shows how this proportion varies by student poverty status, urbanicity, and ACT score threshold used to calculate college-readiness. I label each ACT score with a Michigan postsecondary institution for which the score is the 25th percentile for entering students. EMU and CMU stand for Eastern and Central Michigan University, respectively. UM-Flint and UM-Drbn stand for University of Michigan Flint and Dearborn campuses, respectively.

Figure 1. Supply of College-Ready Students Not Taking a College Entrance Exam.

20, which is the 25th percentile of all students in Michigan in the pre-policy sample who attend and graduate from a four-year postsecondary institution. ACT, Inc. cites

appendix) reports the mean, standard deviation, and various percentiles of the distribution of scores of takers and latent scores of non-takers in the pre-policy period.

a score of 20 as likely qualifying a student for admission to a “traditional” four-year institution.¹⁵ The choice of 20 reflects a threshold that represents students with a good chance of admittance to, and success at, a reasonably selective four-year institution.

In table 2, column 1, I show that 58 percent (117,953 students) of ACT/SAT-takers pre-policy score at or above 20 (row 1). Over 21 percent (26,717) of students not taking either exam would score at this level based on the distribution of latent scores (row 2). This means that if all students took the exam, we would see a 22.7 percent ($= 26,717 \div 117,953$) increase in the number of students scoring college-ready (row 3). Put differently, for every 100 students taking the test and scoring college-ready, there exist another 23 students not taking the test but who would score college-ready. I refer to the fraction 0.227 as the “proportion of college-ready non-takers to takers.” When I consider an ACT score threshold of 22 rather than 20, this proportion decreases somewhat to 0.192.

I calculate standard errors for the proportion of college-ready non-takers to takers, percent of test-takers who score college-ready, and percent of non-takers who score college-ready. I compute these standard errors by running 200 bootstrapped replications of the above exercise and calculating the statistics after each replication. The standard deviation of the statistic across these replications is the estimated standard error of the statistic. The 95 percent confidence interval for the proportion of college-ready non-takers to takers ranges from 0.196 to 0.259.¹⁶

Who Are the High-achieving Non-takers?

I have shown that there is a nontrivial supply of “high-achieving non-takers,” or students who do not take a college entrance exam but would score college-ready if they did. It is important to understand whether this supply varies across different subgroups of the student population. This heterogeneity has implications for which groups of students might experience larger impacts of the mandatory ACT policy on postsecondary outcomes. Moreover, if the supply is larger among disadvantaged populations, this would support explanations for the income gap in college enrollment such as information barriers and complexity in the financial aid and college application process.

In figure 2, I plot the distributions of post-policy scores, pre-policy scores of test-takers, and predicted pre-policy scores of non-takers separately by sex, race, and free lunch status. The first noticeable difference when comparing the frequency distributions of black to white students, or poor to non-poor students, is the far smaller number of disadvantaged students taking a college entrance exam. The second noticeable difference is the lower scores earned by disadvantaged students. As the differences in the supply of college-ready non-takers relative to college-ready takers is difficult to discern visually, I report the results numerically in table 2.

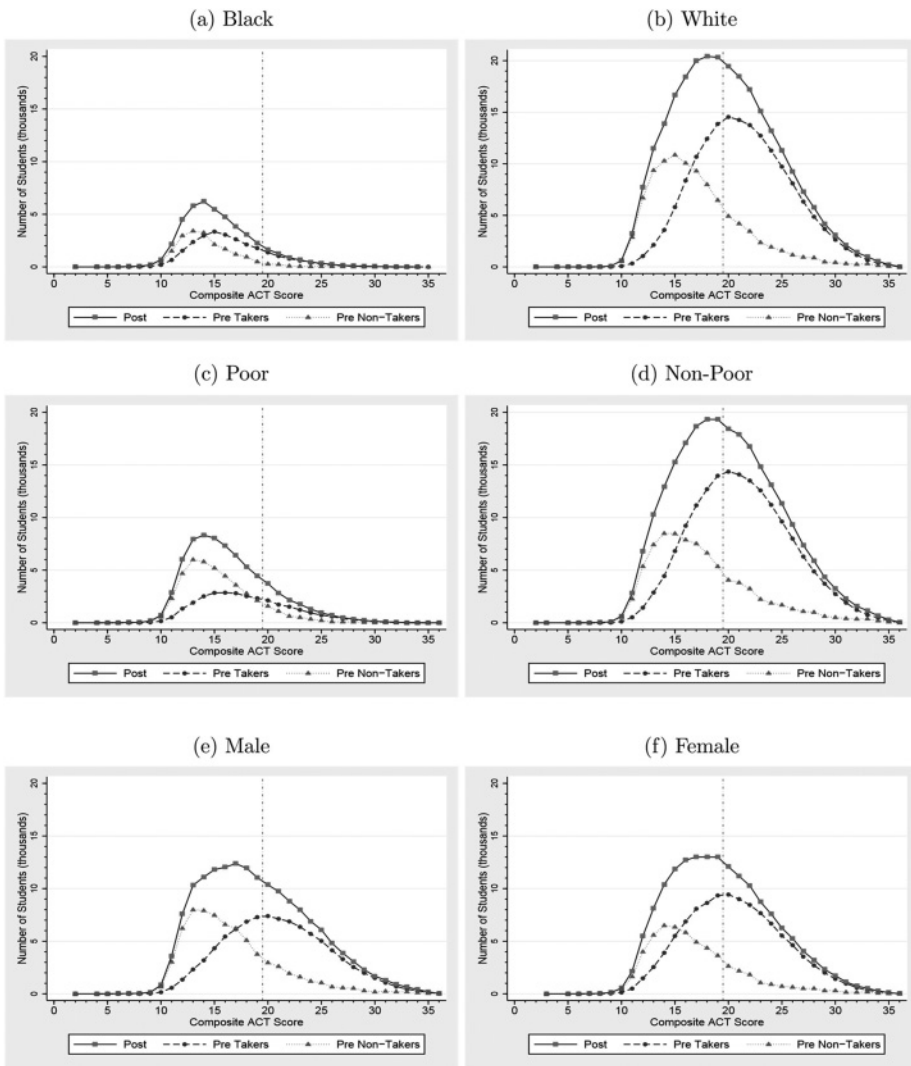
15. See ACT, Inc. (2002). A score of 18–21 likely qualifies a student for admission to nonselective institutions, 20–23 to traditional institutions, 22–27 to selective institutions, and 27–31 (or higher) to highly selective institutions.

16. Each bootstrapped replication resamples entire schools from the original data to allow for correlation of the error term within schools. The main assumption for the validity of the bootstrapped standard errors is that the original sample is representative of the population of interest. This is convincing because the sample is indeed the population of all Michigan public school students, which is the population of interest. See Efron and Tibshirani (1993) for details. Because the standard errors are more conservative, I conduct the bootstrapped replications after having already created the DFL weights using the original sample.

Table 2. Heterogeneity in the Pre-Policy Supply of College-Ready Students Not Taking a College Entrance Exam

	Among Poor Students										
	All (1)	White (2)	Black (3)	Female (4)	Male (5)	Non-Poor (6)	Poor (7)	Urban (8)	Rural (9)	High Gr. 8 Scores (10)	Low Gr. 8 Scores (11)
Scoring college-ready (ACT ≥ 20)											
Percent of ACT/SAT-takers	58.2 (0.9)	64.5 (0.5)	19.4 (2.1)	57.4 (0.9)	59.2 (0.8)	62.5 (0.7)	33.5 (1.2)	20.0 (2.0)	47.0 (0.9)	56.7 (1.0)	16.4 (0.8)
Percent of non-takers (latent score)	21.3 (1.5)	23.9 (1.5)	4.5 (1.4)	21.9 (1.8)	20.8 (1.3)	27.6 (1.8)	11.3 (0.5)	8.8 (1.0)	12.1 (1.0)	32.1 (1.3)	5.4 (0.3)
Proportion of non-takers to takers	0.227 (0.016)	0.221 (0.013)	0.175 (0.060)	0.193 (0.016)	0.265 (0.017)	0.217 (0.014)	0.480 (0.032)	0.544 (0.092)	0.392 (0.043)	0.365 (0.024)	0.651 (0.054)
Scoring college-ready (ACT ≥ 22)											
Percent of ACT/SAT-takers	41.9 (0.8)	47.0 (0.6)	10.1 (1.2)	40.7 (0.9)	43.4 (0.8)	45.7 (0.8)	20.6 (0.8)	10.7 (1.2)	30.3 (0.8)	37.8 (0.9)	7.9 (0.4)
Percent of non-takers (latent score)	13.0 (1.4)	14.6 (1.4)	1.9 (1.0)	13.3 (1.6)	12.8 (1.2)	18.1 (1.8)	5.0 (0.4)	4.5 (0.7)	4.7 (0.8)	16.9 (1.2)	1.5 (0.2)
Proportion of non-takers to takers	0.192 (0.020)	0.185 (0.016)	0.137 (0.065)	0.164 (0.020)	0.222 (0.020)	0.195 (0.018)	0.343 (0.034)	0.518 (0.095)	0.234 (0.049)	0.287 (0.027)	0.383 (0.055)

Notes: The sample is all first-time, public school Michigan eleventh graders in years 2004–09, conditional on reaching spring of eleventh grade. Row 1, column 1 reports that 58.2 percent of students taking the ACT or SAT pre-policy scored at least a 20. Row 2, column 1 reports that 21.3% of students who took neither exam pre-policy would have scored at least a 20. Row 3, column 1 reports that for every 100 students scoring at least a 20 pre-policy, there were 22.7 students who would have scored at least a 20, but who took neither exam. Latent scores of non-takers estimated as explained in text and reweighted to adjust for cohort size and composition following DiNardo, Fortin, and Lemieux (1996). Free lunch status measured as of eleventh grade. Standard errors in parentheses calculated using 200 bootstrap replications.



Notes: Figure shows the distribution of ACT scores post-policy, pre-policy, and the difference, which is the latent score distribution among non-takers in the pre-period. All figures reweight the post-policy sample to resemble the pre-period sample according to observed characteristics following DiNardo, Fortin, and Lemieux (1996). The weights are normed so the pre- and post-policy samples are of equal size. The pre-policy proportion of college-ready non-takers to takers in panels (a), (b), (c), (d), (e), and (f) is 0.175, 0.221, 0.480, 0.217, 0.265, and 0.193, respectively.

Figure 2. Observed and Latent ACT Scores by Subgroup.

The proportion of college-ready non-takers to takers is slightly lower among black students than among white students (table 2, row 3, columns 2 and 3), but the difference is not statistically significant. There is a somewhat larger (and statistically different) supply of male college-ready non-takers relative to female college-ready non-takers (columns 4 and 5). The proportion among men is 0.265, and among women it is 0.193. The most dramatic heterogeneity is seen by poverty status. The proportion of non-poor, college-ready non-takers to takers is near the levels we have seen thus far at 0.217, although the proportion for poor students is 0.480. For every 100 poor students taking

a college entrance exam and scoring at a college-ready level, there are nearly 50 poor students who would score college-ready, but do not take the exam.¹⁷

This large supply of college-ready poor students not taking a college entrance exam provides evidence that the supply of “missing one-offs” identified in recent literature (Bowen, Chingos, and McPherson 2009; Hoxby and Avery 2013; Dillon and Smith 2017) exists earlier in the college application process. Indeed, these high-achieving students have not made it past even the earliest hurdles in the college application process.

Given the large supply of college-ready non-takers among poor students, examining heterogeneity within this group has potential policy relevance. I split poor students by their urban/rural status and by their eighth-grade test score on the state assessment. If promising non-takers are concentrated geographically, this would provide policy makers with a more targeted population at which to aim their reforms. Conditioning on whether students earn high or low eighth-grade test scores is particularly policy-relevant, because teachers and guidance counselors can use these scores to determine their investment of resources during high school.

I find that the proportion of college-ready non-takers to takers is particularly high among poor urban students (0.54), and among poor students with below-average eighth grade test scores (0.65). For every ten such students taking the ACT or SAT and scoring college-ready, there are between five and seven who do not take the exam but would score college-ready. There are smaller but still large populations of these students among poor-rural students (0.39) and among poor students with above-average eighth grade test scores (0.37). These results suggest that teachers and guidance counselors should not assume that disadvantaged students who score poorly on state assessments would not be qualified to enroll in a four-year college, if set on the proper path.¹⁸

In panel B of figure 1, I examine the sensitivity of the results to the choice of college-readiness threshold. The x -axis is the ACT score used as the threshold. I additionally label most ACT scores on the x -axis with a Michigan postsecondary institution for which the score is the 25th percentile for entering students. The y -axis gives the proportion of college-ready non-takers to takers.

Figure 1b reveals two interesting points: First, whereas the proportion of college-ready non-takers to takers among the overall sample is relatively stable across the choice of college-readiness threshold (solid line), the proportion among poor students (dashed line) varies greatly depending on the choice of threshold. Lowering the threshold from 20 to 18 increases the proportion to about two-thirds, and raising the threshold from 20 to 22 decreases the proportion to just over one-third.

Second, when we look by urban/rural status among low-income students, the proportion of urban college-ready non-takers to takers remains quite high across ACT score thresholds in the mid 20s. Alternatively, the proportion among low-income rural students drops steeply as the threshold increases. This result suggests that high-achieving,

17. For each calculation by subgroup, I restrict the sample to students in that group and create a new set of DFL weights scaled to adjust for the different sample sizes pre- versus post-policy. Thus, the larger number of free-lunch eligible or minority students in the post period does not mechanically lead to a larger proportion of college-ready non-takers to takers among these groups.

18. I also examine heterogeneity by school-level characteristics, such as school poverty share, and whether the school was a pre-policy ACT test center—two subgroups that are of interest later in the paper. I find a slightly larger proportion among high poverty high schools and among schools that were not a pre-policy center, but the differences across the groups are not statistically significant.

low-income students not embarking on the path toward four-year college enrollment by taking a college entrance exam occurs less prominently in rural than in urban areas. Hoxby and Avery (2013), on the other hand, find a large supply of high-achieving, low-income students in rural areas who score very well on college entrance exams but do not apply to selective colleges. Taken together, our results suggest that high-achieving, low-income students in rural areas tend to get far enough in the college application process to take a college entrance exam, but then “undermatch” in their application behavior. I find that many such students in urban areas fail to even get to the point of taking a college entrance exam.

I present several robustness checks and supplementary analyses in the online Appendix B. First, I show that the results are not sensitive to noncompliance in the post-period (i.e., the fact that fewer than 100 percent of students take the ACT). Second, I conduct a similar analysis, and find similar results, in several other mandatory ACT states, showing that the results in Michigan are generalizable. Finally, I conduct the Michigan analysis redefining “college-ready” as having an ACT score of at least 20 and having a high school grade point average above some threshold. I find that although the proportion of college-ready non-takers falls after conditioning on high school grade point average, the proportion remains substantial, especially for poorer students.

5. EFFECTS OF THE MANDATORY ACT POLICY ON POSTSECONDARY OUTCOMES

Effects on College Enrollment and Choice

The simplest way to examine the effect of the mandatory ACT policy on college enrollment is to compare enrollment before and after the policy. As previously shown in table 1, the average postsecondary enrollment rate among the three pre-policy cohorts in my sample is 0.570. The average rate among the three post-policy cohorts is 0.589, or 1.9 percentage points higher. The increase in the enrollment rate at four-year colleges is 1.0 percentage points. Controlling for student-level demographics and eighth grade scores, as well as school fixed effects, decreases the overall pre/post difference in college enrollment from 1.9 to 1.4 percentage points, and increases the four-year enrollment difference from 1.0 to 1.1 percentage points.

These increases may not represent the true impact of the mandatory ACT policy. The sinking economy, shifting demographic composition, similarly timed education reforms, and any other factors changing over this time period, could affect the college enrollment of Michigan students.¹⁹

To mitigate the biases resulting from these omitted factors, I estimate the causal impact of mandatory ACT-taking on postsecondary enrollment in Michigan using a DID research design. Specifically, I compare changes in college attendance between the pre- and post-policy periods in schools that did not have an ACT test center in the

19. As examples of similarly timed education reforms, the Michigan Promise Scholarship was a short-lived merit scholarship that offered up to \$4,000 toward college for the last pre-mandatory ACT policy cohort and first two post-policy cohorts. Preliminary findings suggest little to no impact of the policy on college-going (Dynarski et al. 2013). The Michigan Merit Curriculum, also implemented around this time, increased the course requirements necessary to graduate high school. The first cohort exposed to the policy was in eleventh grade in 2010, however, and thus not in my sample.

school pre-policy, to those that did. I estimate the following equation using OLS:

$$Y_{isdt} = \beta_0 + \beta_1 Post_t + \beta_2 NoCenter_{isd} + \beta_3 (Post_t * NoCenter_{isd}) + \beta_4 X_{isdt} + \alpha_s + \varepsilon_{isdt}, \quad (2)$$

where Y_{isdt} is a postsecondary outcome for student i in school s in district d in cohort t . $Post$ is a dummy for attending eleventh grade post-policy, $NoCenter$ is a dummy for attending a school without a pre-policy ACT test center (which drops out when I include school fixed effects), X is a vector of student-level and school- and district-year level covariates, and α is a full set of school fixed effects.²⁰ ε is the error term clustered at the school level. β_3 , the coefficient of interest, is the effect of the policy in schools with no pre-policy test center relative to those with a center.

The intuition behind the above strategy is that schools without a test center will experience a slightly larger increase in ACT-taking because of the mandatory ACT policy than will schools with a pre-existing test center. The identifying assumption behind my estimation strategy is that any differential changes in college enrollment after the mandatory ACT policy between the students in these two groups of schools are due to the effects of the policy. Other similarly timed statewide education reforms or factors that are changing over time, and could affect college-going, are assumed to affect the two types of schools equally.

Columns 1 and 2 of table 3 show student-weighted sample means of schools with and without a test center before the mandatory ACT policy. Slightly over half of students attend a school with a test center, even though there are double the number of schools without a center. Not only are schools with test centers much larger, but they tend to enroll students with higher academic achievement, higher ACT-taking rates, and higher educational attainment. Schools with a test center are more likely to be in an urban or suburban area, and less likely to be in a rural area.²¹

Given the DID design, the threat to validity is not if the two types of schools are different but rather if they are changing differentially over time. In columns 4 and 5 of table 3, I show means at the two types of schools in the post period, and the DID estimate in column 7. There is some evidence that the populations of these schools are changing differentially over time. There is an increase in free lunch status for schools without a center over time, relative to schools with a center, and a decrease in eleventh grade enrollment.

To ensure that the schools with and without a test center are similar except for their test center status, I use propensity score matching on a series of school- and district-year level observed characteristics to create a sample of matched test center and

20. Unless otherwise noted, X includes student-level sex, race, free lunch status, LEP, SPED, and eighth grade test score; school-year level fraction black, fraction free-lunch eligible, number of eleventh graders and mean eighth grade scores; and the same district-year level covariates plus guidance counselor–pupil ratio, dummies indicating urban/rural status, and the local unemployment rate.

21. It is not surprising that schools with a center are quite different than those without, as becoming a test center is primarily a demand-driven phenomenon. To become a test center, a teacher, counselor, or administrator from the school fills out an online form. They agree to be open on at least one testing day per year, must expect at least 35 students on the testing day, and must have the proper room conditions and seating arrangements, which are then verified by an ACT official.

Table 3. Sample Means Pre- and Post-Policy by Pre-Policy Test Center Status

	Before Mandatory ACT Policy			After Mandatory ACT Policy			Diff-in-Diff (6) - (3)	Matched Sample Diff-in-Diff (8)
	No Center (1)	Center (2)	Difference (3)	No Center (4)	Center (5)	Difference (6)		
Demographics								
Black	0.124	0.166	-0.043*	0.145	0.180	-0.035	0.008	0.003
Hispanic	0.032	0.029	0.003	0.037	0.032	0.005	0.002	0.000
Free lunch	0.248	0.220	0.028*	0.331	0.292	0.040**	0.012*	-0.006
Eighth grade scores	-0.009	0.071	-0.080**	-0.025	0.056	-0.080**	0.000	-0.010
Pupil-Teacher ratio	20.6	21.8	-1.2	19.8	20.1	-0.3	0.9	1.5
Grade 11 enrollment	216.6	345.1	-128.5***	223.3	360.1	-136.8***	-8.3*	-6.7
Local unemployment	7.57	7.11	0.45*	9.26	8.83	0.43	-0.024	-0.057
Urban area	0.543	0.711	-0.167***	0.551	0.714	-0.163***	0.004	0.005
Rural area	0.457	0.289	0.167***	0.449	0.286	0.163***	-0.004	-0.005
Educational attainment								
Take ACT or SAT	0.540	0.607	-0.067***	0.927	0.932	-0.005	0.061***	0.039***
Graduate high school	0.847	0.876	-0.029***	0.847	0.879	-0.032***	-0.003	-0.001
Enroll in any college	0.554	0.611	-0.056***	0.576	0.631	-0.055***	0.001	0.005
Enroll in four-year college	0.292	0.343	-0.050***	0.306	0.352	-0.046***	0.004	0.008*
Enroll in two-year college	0.262	0.268	-0.006	0.270	0.279	-0.009	-0.003	-0.003
Number of schools	523	251		518	251			
Number of students	165,009	181,463		168,825	186,468			

Notes: The sample is all first-time, public school Michigan eleventh graders in years 2004–09, conditional on reaching spring of eleventh grade. “No Center” and “Center” refer to whether or not a high school was an ACT test center before the mandatory ACT policy. The sample for column 8 is restricted to the 226 schools without a pre-policy ACT test center and the 226 schools with a pre-policy test center matched using nearest neighbor matching.

*Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

non-test-center schools.²² I use nearest neighbor matching (without replacement), because it tends to produce the best balance of covariates in my sample. I show that my results are not sensitive to either propensity score reweighting, or to other methods of matching such as kernel or caliper matching, that have been shown to produce superior results in some contexts (Heckman, Ichimura, and Todd 1997; Busso, DiNardo, and McCrary 2013). Because some of the schools with a test center have extremely high propensity scores where there are few similar non-test center schools, I trim the ten percent of schools with the highest propensity scores—these tend to be very large schools in suburban areas. Trimming fewer of the center-schools produces similar results but inferior covariate balance.²³

I find that after the propensity score matching, there is no evidence that schools with and without a test center are trending differentially with respect to their composition

22. The following covariates are included in the propensity score regression: (1) school- and district-level pupil-teacher ratio, percent free-lunch eligible, grade eleven enrollment, and fraction black; (2) average school-level eighth and eleventh grade test scores; (3) dummies for school urban/rural status; (4) the growth rate in the school’s eleventh grade enrollment; (5) the district-year level guidance counselor-pupil ratio; and (6) the local unemployment rate.
23. If I trim the sample by 20 percent, my college enrollment results display the same pattern of heterogeneity and are slightly larger in magnitude. If I do not trim any of the test center schools with the highest propensity scores, the balance of covariates across the two types of schools is substantially worse and the pattern of heterogeneity is again the same, but slightly smaller in magnitude.

(see column 8, table 3). None of the covariates has a statistically significant DID estimate. Rates of ACT-taking at schools without a pre-policy center nonetheless increase by 4 percentage points after the policy relative to schools with a pre-policy center. This 4-percentage-point gap arguably captures the effect on test-taking of having a test center in one's high school. There is no DID effect on high school graduation or overall college enrollment, but a marginally statistically significant 0.8 percentage point increase in four-year enrollment.

It is important to note that most stories involving differences in unobservables biasing the effects would provide a downward bias on the results. For example, if particularly active or motivated teachers, counselors, or administrators are those who initiate a test center at a school, it seems likely that such staff would more effectively implement the mandatory ACT policy or engage in other practices aimed at boosting enrollment than staff at non-test-center schools.

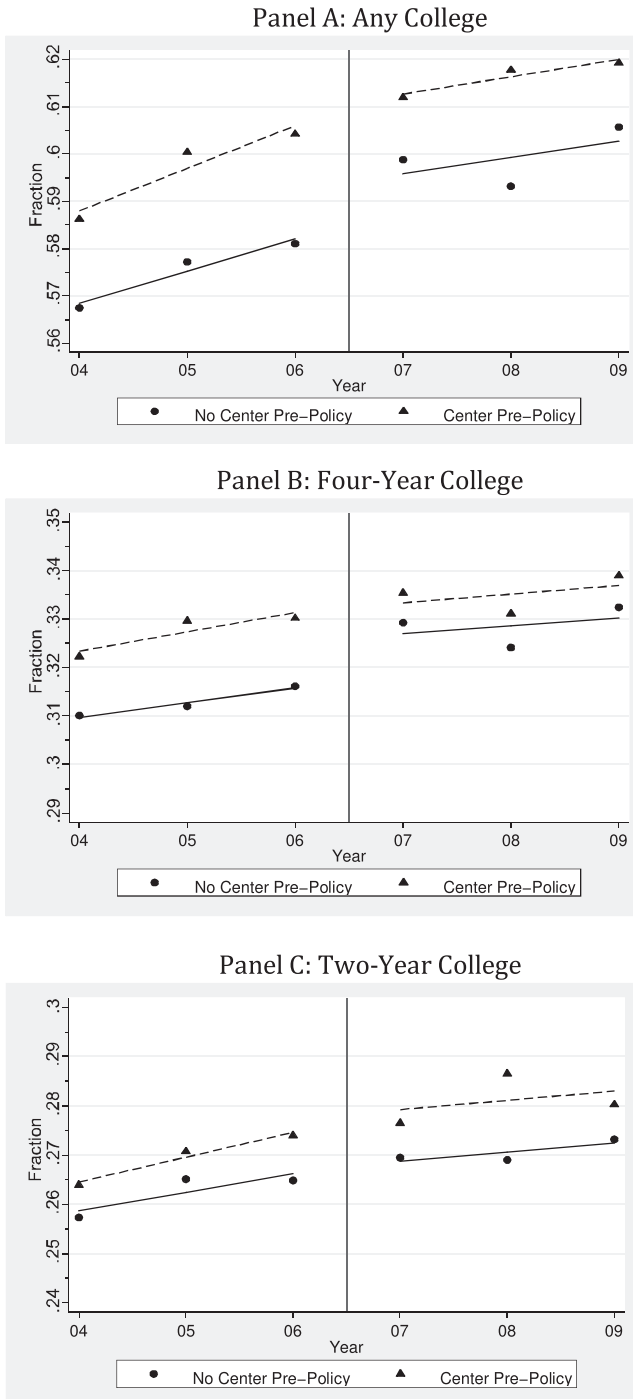
To further test the validity of the DID methodology, I plot college attendance rates of schools in the matched sample by cohort and test center status. Trends in college enrollment are nearly identical across the two types of schools prior to the mandatory ACT policy (figure 3). This suggests that college enrollment would have continued to trend in parallel in the absence of the policy, satisfying one of the key identifying assumptions of my estimation strategy. The pre-policy level of four-year college enrollment is higher in the matched sample of schools with a test center, presumably reflecting that some of the students induced into taking the ACT by having a center in their school subsequently enroll in a four-year college.

The regression-adjusted DID results estimated using equation 2 show little effect of the policy on overall enrollment regardless of covariates, school fixed effects, or matching method (table 4, row 1, columns 1–5). The point estimate is between 0.3 and 0.5 percentage points, is statistically insignificant, and is fairly stable across the columns. The effect on the probability that a student enrolls at a four-year institution is 0.8 percentage points (standard error of 0.4 percentage points—column 6).²⁴ Panel B of figure 3 depicts this DID effect visually. Adding covariates does not alter the estimate but the inclusion of school fixed effects lowers the coefficient to 0.6 percentage points. This represents a 1.9 percent increase in the four-year enrollment rate, off of the pre-policy mean of 32.1 percent. There is a smaller corresponding negative (and statistically insignificant) point estimate for two-year enrollment.²⁵

The coefficient on the *Post* dummy in column 8 of table 4 indicates a 1.1 percentage point increase in four-year enrollment post policy among students at schools with a test center pre-policy. The 0.6 percentage point increase for the non-test center schools is above and beyond this increase. Although the 1.1-point increase may in part be driven by the policy change, I cannot disentangle the effects of the policy for schools with a pre-policy center from other factors changing over time. In this sense, the DID effect that I estimate likely represents a lower bound of the policy's impact.

24. Note that the standard errors do not account for the propensity score matching. Eichler and Lechner (2002) show that in their sample the standard errors that ignore the matching are similar to bootstrapped standard errors that take the matching into account.

25. I define two-year enrollment as enrolling in a two-year school and not a four-year school, so that two- and four-year enrollment are mutually exclusive. Estimates of the effect of the policy on enrollment at selective four-year or out-of-state colleges were statistically imprecise.



Notes: Figure shows college enrollment pre- and post-mandatory ACT by whether or not a student attends a school with an ACT test center pre-mandatory ACT. The sample is restricted to the propensity score matched sample of high schools. Panel A includes any college enrollment, Panel B includes four-year enrollment only, and Panel C includes two-year enrollment.

Figure 3. College Enrollment by Cohort and Pre-Policy Test Center Status.

Table 4. The Effect of the Mandatory ACT on Postsecondary Enrollment

	Dep. Var. = Any Enrollment					Dep. Var. = Four-Year Enrollment					Dep. Var. = Two-Year Enrollment				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Nearest Neighbor Matching	Nearest Neighbor Matching	Kernel Matching	Kernel Matching	P-Score Weighting	Nearest Neighbor Matching	Nearest Neighbor Matching	Kernel Matching	Kernel Matching	P-Score Weighting	Nearest Neighbor Matching	Nearest Neighbor Matching	Kernel Matching	Kernel Matching	P-Score Weighting
Post * No test center in school	0.005 (0.005)	0.003 (0.004)	0.003 (0.004)	0.003 (0.004)	0.004 (0.004)	0.008* (0.004)	0.008* (0.004)	0.006 (0.004)	0.006 (0.004)	0.007* (0.004)	-0.003 (0.004)	-0.005 (0.005)	-0.003 (0.004)	-0.002 (0.004)	-0.002 (0.004)
Post	0.019*** (0.003)	0.029*** (0.005)	0.016*** (0.003)	0.015*** (0.003)	0.014*** (0.003)	0.008** (0.003)	0.016*** (0.004)	0.011*** (0.003)	0.011*** (0.003)	0.011*** (0.003)	0.012*** (0.003)	0.013*** (0.005)	0.005 (0.003)	0.004 (0.003)	0.004 (0.003)
No test center in school	-0.022 (0.015)	-0.014 (0.011)				-0.015 (0.015)	-0.014* (0.008)				-0.007 (0.012)	0.000 (0.011)			
Covariates	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
School fixed effects	N	N	Y	Y	Y	N	N	Y	Y	Y	N	N	Y	Y	Y
Pre-policy mean		0.587		0.590	0.588		0.321		0.320	0.317		0.266		0.270	0.271
Sample size		536,813		614,974	701,765		536,813		614,974	701,765		536,813		614,974	701,765

Notes: The sample is all first-time, public school Michigan eleventh graders in years 2004–09, conditional on reaching spring of eleventh grade. For columns 1–3, 6–8, and 11–13, the sample is restricted to the 226 schools without a pre-policy ACT test center and the 226 schools with a pre-policy test center matched using single nearest neighbor matching without replacement. An Epanechnikov kernel and bandwidth of 0.06 is used in columns 4, 9, and 14. Each column is a separate linear probability model regression. Standard errors in parentheses are clustered at the school level.

* Significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Heterogeneity of Impacts

It seems unlikely that all students would be equally impacted by the mandatory ACT policy. Many students would take the ACT regardless of the policy. Other students are forced to take the ACT, but are so academically unprepared—or otherwise off the path of application to college—that being forced to take the exam will have no impact on their educational plans. In this section I estimate heterogeneity in the effects of the policy on college-going. This heterogeneity captures differences across groups both in treatment dosage (i.e., some groups will experience larger effects on ACT-taking) and in sensitivity of college-going to a given dosage.

To home in on the marginal student most impacted by this policy, I create an index measuring the predicted probability that a student would take the ACT based on the pre-policy relationship between ACT-taking and student-level observed demographic characteristics. Specifically, I estimate the following equation using OLS:

$$TAKE_{isdt} = \beta_0 + \beta_1 X_{isdt} + \alpha_s + \varepsilon_{isdt}, \quad (3)$$

where X includes all main effects and interactions of sex, race, free and reduced-price lunch status, and LEP and SPED status. α is again a full set of school fixed effects.²⁶ I estimate this equation using only pre-policy students, then predict \widehat{TAKE} for all students pre- and post-policy, thus creating for all students a predicted probability of taking the ACT in the absence of the policy.²⁷

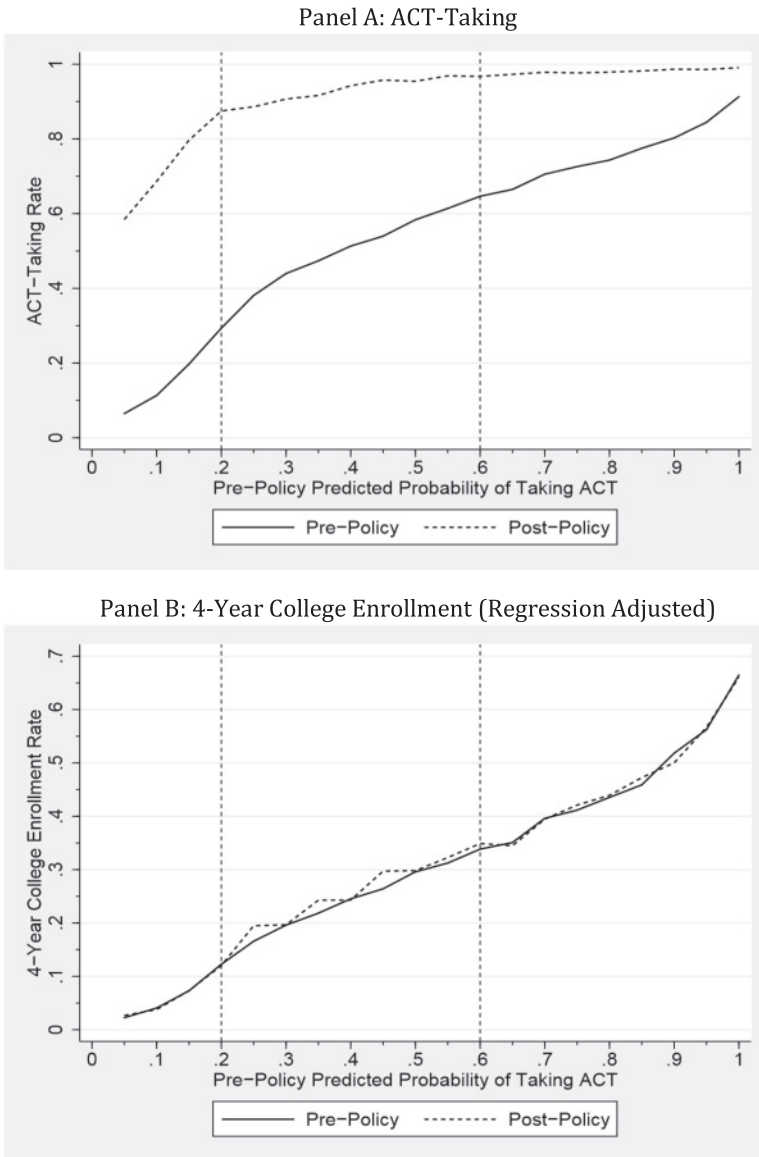
I show that the mandatory ACT policy increases ACT-taking most for students with the lowest predicted probability. Panel A of figure 4 breaks students into vigintiles (twenty quantiles) based on this index, and plots mean ACT-taking rates of students in pre-policy cohorts (solid line) and of students in post-policy cohorts (dashed line). The distance between the two lines in this figure represents the treatment dosage, in the sense that it gives the change in the ACT-taking rate for students with a given probability of taking the ACT pre-policy. Table 5 reports the DID effects of the policy on ACT-taking and college enrollment for all students, and by quintiles of this predicted probability index. Among all students, there is a 3.4 percentage point effect of the policy on ACT-taking in non-test center high schools, relative to test center schools (column 1, row 1). The increases are largest for students with the lowest pre-policy probability (row 1, columns 2–6), with no change for high-probability students.

The remaining rows of the first column in table 5 replicate the preferred specification from Table 4. Despite the large impact on ACT-taking among students with a very low pre-policy probability, the effects on four-year enrollment are near zero for this group, as they are for students in the top two quintiles of the probability index. Effects are largest on four-year college enrollment for students with a low or mid-level probability.²⁸ In Panel B of figure 4, I plot the pre-policy raw four-year college enrollment rates for each vigintile of the predicted probability of ACT-taking (solid line). I then estimate

26. Appendix table B.2 (available online) reports the results from this regression. The results are nearly identical when using probit or logit.

27. Abadie, Chingos, and West (2012) show that forming subgroups based on a predicted outcome fitted within the control group can cause biases. This is not the case here due to my use of the difference-in-differences estimator as opposed to a simple comparison of the outcome in the pre-versus post-policy period. The difference in the fit of the prediction between the pre- and post-policy students will not vary differentially across schools with and without a pre-policy test center.

28. Results are similar when dividing the predicted probability index by tercile or quartile.



Notes: Panel A plots the ACT-taking rate pre- and post-mandatory ACT at twenty quantiles of the predicted probability that a student would take the ACT based on the pre-policy relationship between observed characteristics and ACT-taking. Panel B plots the raw, pre-policy four-year enrollment rate among students in the matched sample of high schools (solid line) at these same twenty quantiles. It then adds to this line the difference-in-difference four-year enrollment effect of the policy (dashed line). Note the smaller scale of the y-axis in Panel B to more clearly show the difference between the two lines.

Figure 4. ACT-Taking and College Enrollment by Predicted Probability of ACT-Taking.

equation 2 separately for each vigintile and add the DID coefficient to the pre-policy rate (dashed line). As seen in table 5, the enrollment effects are entirely concentrated within the second and third quantiles of the predicted probability index.

To increase precision and collapse students into a group that seems marginal, and a group whose college enrollment behavior seems relatively unaffected by the policy, I

Table 5. Using Students' Predicted Probability of ACT-Taking Pre-Policy to Narrow in on the Marginal Student

Dependent Variable	Pre-Policy Probability (Take ACT)							
	All (1)	Very Low (2)	Low (3)	Middle (4)	High (5)	Very High (6)	Low/Middle (7)	Tails (8)
Take ACT	0.034*** (0.013)	0.044*** (0.012)	0.038*** (0.010)	0.028*** (0.006)	0.007 (0.007)	0.007 (0.012)	0.032*** (0.006)	0.036** (0.018)
	<i>0.580</i>	<i>0.199</i>	<i>0.457</i>	<i>0.600</i>	<i>0.710</i>	<i>0.835</i>	<i>0.531</i>	<i>0.618</i>
Enroll in:								
Any college	0.003 (0.004)	-0.001 (0.008)	0.013 (0.008)	0.014** (0.007)	-0.008 (0.007)	0.003 (0.007)	0.014** (0.006)	-0.003 (0.004)
	<i>0.587</i>	<i>0.305</i>	<i>0.497</i>	<i>0.616</i>	<i>0.676</i>	<i>0.765</i>	<i>0.559</i>	<i>0.608</i>
Four-year college	0.006 (0.004)	-0.002 (0.005)	0.013** (0.007)	0.012** (0.006)	0.001 (0.008)	0.001 (0.008)	0.013** (0.005)	0.000 (0.004)
	<i>0.321</i>	<i>0.077</i>	<i>0.207</i>	<i>0.305</i>	<i>0.398</i>	<i>0.553</i>	<i>0.259</i>	<i>0.369</i>
Two-year college	-0.003 (0.004)	0.001 (0.007)	-0.000 (0.007)	0.001 (0.007)	-0.010 (0.007)	0.002 (0.007)	0.001 (0.005)	-0.003 (0.004)
	<i>0.266</i>	<i>0.227</i>	<i>0.290</i>	<i>0.311</i>	<i>0.277</i>	<i>0.212</i>	<i>0.301</i>	<i>0.239</i>
Covariates	Y	Y	Y	Y	Y	Y	Y	Y
School fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Sample size	536,813	86,136	117,944	117,381	104,082	111,270	235,325	301,488

Notes: The sample is all first-time, public school Michigan eleventh graders in years 2004–09, conditional on reaching spring of eleventh grade. The sample is restricted to the 226 schools without a pre-policy ACT test center and the 226 schools with a pre-policy test center matched using nearest neighbor matching. Each point estimate is from a separate linear probability model, difference-in-difference regression. Standard errors in parentheses are clustered at the school level. Pre-policy dependent variable means are in italics below the standard errors.

** Significant at the 5% level; *** significant at the 1% level.

combine the low and middle students together, and the very low, high, and very high students together. I call this latter group the “tails” of the distribution, capturing students who either would have taken the ACT regardless or who are so off the college track that taking it makes no difference for their college-going behavior. Among students in the low to middle range of the predicted probability index (between the two vertical lines in figure 4), there is a 1.3 percentage point, or 5 percent, increase in enrollment at four-year colleges. There is no effect among students in the tails of the distribution, and the difference across groups is statistically significant (p -value = 0.05).

To guide policy, it would also be helpful to examine which types of students along specific observed dimensions have college enrollment behavior that is most influenced by the mandatory ACT. Table 6 presents results separately by race, sex, and poverty status. Although the effects among black students are imprecisely estimated, boys and poor students (those eligible for free lunch) appear to experience relatively large gains of approximately 1 percentage point. These gains represent a near 3.5 percent increase for boys and a 6 percent increase for poor students relative to their pre-policy mean, and both point estimates are statistically significant at the 5 percent level. Unfortunately, the estimates are not precise enough to reject equality across groups.

Finally, I examine the effects by school poverty share. This is a particularly policy-relevant dimension, as education policies are easier to implement at the school level than only to students with particular characteristics. I split students into terciles based on the share of students in their school who qualify for free or reduced-price lunch. I then combine students in the low- and middle-poverty schools, and compare the

Table 6. Heterogeneity in the Effect of the Mandatory ACT by Student Demographics and School Poverty Share

Dependent Variable	School Poverty Share								
	All (1)	White (2)	Black (3)	Female (4)	Male (5)	Non-Poor (6)	Poor (7)	Low/Middle (8)	High (9)
Enroll in:									
Any college	0.003 (0.004)	0.003 (0.004)	0.003 (0.011)	-0.000 (0.005)	0.005 (0.005)	-0.001 (0.004)	0.016** (0.007)	-0.000 (0.004)	0.009 (0.007)
	<i>0.587</i>	<i>0.605</i>	<i>0.515</i>	<i>0.622</i>	<i>0.552</i>	<i>0.640</i>	<i>0.415</i>	<i>0.634</i>	<i>0.494</i>
Four-year college	0.006 (0.004)	0.005 (0.004)	0.009 (0.009)	0.002 (0.005)	0.009** (0.004)	0.004 (0.004)	0.010** (0.005)	0.001 (0.004)	0.013** (0.006)
	<i>0.321</i>	<i>0.334</i>	<i>0.256</i>	<i>0.350</i>	<i>0.291</i>	<i>0.370</i>	<i>0.164</i>	<i>0.368</i>	<i>0.228</i>
Two-year college	-0.003 (0.004)	-0.002 (0.004)	-0.006 (0.009)	-0.002 (0.005)	-0.004 (0.004)	-0.005 (0.004)	0.006 (0.006)	-0.001 (0.004)	-0.004 (0.006)
	<i>0.266</i>	<i>0.271</i>	<i>0.259</i>	<i>0.272</i>	<i>0.261</i>	<i>0.271</i>	<i>0.251</i>	<i>0.266</i>	<i>0.267</i>
Covariates	Y	Y	Y	Y	Y	Y	Y	Y	Y
School fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample size	536,813	417,851	83,061	268,573	268,240	384,331	148,147	358,113	178,700

Notes: The sample is as in Table 5. Each point estimate is from a separate linear probability model, difference-in-difference regression. Free lunch is measured as of eleventh grade. Standard errors in parentheses are clustered at the school level. Pre-policy dependent variable means are in italics below the standard errors.

**Significant at the 5% level.

effects on those in high-poverty schools. Students in high-poverty schools experience a statistically significant increase in four-year enrollment of 1.3 percentage points or 5.7 percent (table 6, column 9). There is no impact among students at schools with low to middle levels of poverty, and the p -value for the test of equality across the two groups is 0.11.²⁹

Do Marginal Enrollees Drop Out?

Although college entry has been rising in recent decades, college completion has remained flat (Bound, Lovenheim, and Turner 2010). A key concern with a policy such as the mandatory ACT is that it may induce marginal students to attend but not persist through college. If this is the case, then the effects on four-year enrollment rates would overstate the benefits of the program.

In table 7, I present the effects of the policy on the share of students who enroll in a four-year college and persist to the second, third, and fourth years. If all students induced into college by the policy subsequently dropped out, then these point estimates would equal zero. As a reminder, the definition of enrollment is whether a student enrolls by the second fall following on-time high school graduation. Given that my data capture enrollment through summer 2013, students in the most recent cohort who enrolled in college during the second fall after on-time high school graduation have only had time to progress through their second year of college. Consequently, this exercise requires dropping one or more post-policy cohorts from the sample. Row 1, column 1,

29. To further explore effect heterogeneity, in Appendix table A.3 (available online), I present results by eighth-grade test score, which proxies for student ability. I find that the effects are driven by both low- and high-ability students.

Table 7. Examining Whether Four-Year Enrollment Effects Persist

Dependent Variable	Three Pre-Policy Cohorts Plus:			Pre-Policy Dep. Var. Mn. (4)
	All 3 Post Cohorts (1)	First 2 Post Cohorts (2)	First Post Cohort Only (3)	
Enroll within two years	0.006 (0.004)	0.007* (0.004)	0.007 (0.005)	0.321
and persist to year 2	0.004 (0.003)	0.005 (0.004)	0.006 (0.004)	0.278
and persist to year 3		0.004 (0.003)	0.006 (0.004)	0.259
and persist to year 4			0.007* (0.004)	0.244
and graduate in four years			0.005* (0.003)	0.096
Enroll within one year	0.006* (0.003)	0.006* (0.004)	0.006 (0.004)	0.291
and persist to year 2	0.005 (0.003)	0.005 (0.003)	0.006 (0.004)	0.256
and persist to year 3	0.004 (0.003)	0.005 (0.003)	0.006 (0.004)	0.240
and persist to year 4		0.005 (0.003)	0.006* (0.004)	0.228
and graduate in four years		0.002 (0.002)	0.004 (0.003)	0.091
and graduate in five years			0.004 (0.003)	0.169
Sample size	536,813	448,234	357,181	
Covariates	Y	Y	Y	
School fixed effects	Y	Y	Y	

Notes: The sample is as in tables 5 and 6. Each point estimate is from a separate linear probability model, difference-in-difference regression. Standard errors in parentheses are clustered at the school level.

*Significant at the 10% level.

reports the previously estimated four-year enrollment result for the full sample. Columns 2 and 3 show the effect of dropping the most recent and two most recent post-policy cohorts, respectively, each yielding a point estimate of 0.07.

The second row shows the effect on enrolling and persisting to the second year. Among the full sample, the effect is somewhat attenuated to 0.4 percentage points. The effect in percent terms shrinks from 1.9 percent to 1.4 percent given the smaller pre-policy fraction of students enrolling and persisting to the second year (column 4). Examining the effect of the mandatory ACT policy on persisting to the third and fourth years of college requires dropping post-policy cohorts from the sample. The effect on enrolling and persisting to the third year is again 0.4 percentage points (row 3, column 2), but on persisting to the fourth year is 0.7 percentage points (row 4, column 3; significant at the 10 percent level), and the same as the effect on enrolling for that sample. Although the results are imprecise and vary by sample and persistence measure, it appears that students induced to enroll by the policy persist through college at a similar rate as inframarginal students. At the very least, I can reject with 90 percent confidence that all students induced to enroll drop out by their fourth year of college.

The implementation of the policy is too recent to accurately assess if there are increases in degree completion, but I attempt to take a first glimpse at this important measure. The effect on enrolling and then earning a bachelor's degree within four years is a statistically significant 0.5 percentage points (row 5, column 3), or 5.2 percent. I also examine effects on degree receipt within five years. Doing so, however, requires that I redefine the enrollment measure to include only those enrolling by the first fall following on-time high school graduation. The enrollment effect using this measure (0.6 percentage points) is the same as before and marginally statistically significant. The bottom row of table 7 shows that the effect on five-year degree receipt is 0.4 percentage points, or 2.4 percent compared to the 2.1 percent effect on enrollment. The results are imprecisely estimated, but suggest that students induced to enroll by the policy earn a degree at a similar rate as inframarginal students. These results are consistent with other recent studies showing that students induced into colleges by dismantling barriers to the college application process persist at high rates (Bettinger et al. 2012; Carrell and Sacerdote forthcoming).

Robustness Checks

In this section, I briefly describe and summarize results from several robustness checks that examine the sensitivity of my estimates. In the online Appendix B, I discuss the details of these analyses and present complete results (see Appendix table B.1). The first check estimates the DID equation controlling for pre-trending of the outcome variable. Given the relatively few data points (three) before the policy change over which to estimate the pre-trend, this is not my preferred specification. Nevertheless, the results controlling for the pre-trends are slightly attenuated, but very similar to the main results.

The second robustness check uses a different method of constructing the treatment and comparison groups. Instead of grouping students by their high schools' pre-policy test center status, I use a student's home address during the eleventh grade, and the address of the nearest pre-policy test center, to group students by whether they live far from (treatment) or close to (comparison) the nearest pre-policy center. This strategy serves as a test of the external validity of the matched sample to the entire Michigan sample, as well as a test of the sensitivity of the results to the different method of constructing the treatment/comparison group.³⁰ Among the propensity score matched sample of schools, the effects of the policy on postsecondary outcomes are similar using the distance measure and show the same pattern of heterogeneity, with coefficients that are generally greater in magnitude and more precisely estimated. The results and pattern of heterogeneity are still similar when not restricting the analysis to the matched sample of schools, suggesting the effects of the policy can be extrapolated to the entire population of Michigan.

30. I prefer the school-level test center method as my main strategy, and the distance method as a robustness check for two reasons: (1) separating students by distance into treatment and comparison groups is arbitrary because distance is a continuous measure, and (2) it is easier to understand the selection process of schools becoming test centers than of students living close to or far from a test center. Thus, I can more convincingly sign any possible bias due to selection on unobserved characteristics when using the test center strategy than when using the distance strategy.

6. DISCUSSION

Interpretation of Effects

The effects estimated in this paper using the DID design may represent a lower bound of the statewide policy impact. There is likely some portion of the effect that is not captured by this methodology because it is experienced equally by students at schools with a pre-policy test center and those without. Another way to characterize the effects is that they are local average treatment effects (LATEs) estimated for a specific and marginal group of students. The LATE is the expected outcome gain for those induced to receive treatment through a change in the instrument (Imbens and Angrist 1994). In this context, these are post-policy ACT-takers who were enrolled in a high school without a pre-policy center and would not have taken a college entrance exam pre-policy in their high school, but would have if enrolled at a high school with a center.

To obtain a treatment on the treated estimate for this group of students, I scale the effects on four-year enrollment by the first-stage DID increase in ACT-taking. Doing so yields a treatment on the treated estimate suggesting that 18 percent of this marginal group of students would subsequently enroll in a four-year college ($= 0.6 / 3.4$).³¹ This result is consistent with the large treatment effects often realized by marginal students picked up by LATEs in the context of education policies (Card 1995). If the results were scalable, however, we would expect to see statewide increases in four-year enrollment rates of 18 percent as a result of the policy.

This number represents one possible upper bound of the policy's impact, yet it seems extraordinarily high. Hurwitz et al. (2015) estimate the effect of a mandatory SAT policy in Maine using a DID approach. They estimate that the policy increased the four-year enrollment rate from 4 to 6 percent. This magnitude of effects is far closer to the main effect of the policy that I estimate (2 percent) than the 18 percent upper bound calculated above.

Capacity Constraints

Another issue regarding the interpretation of my results involves supply-side capacity constraints on the side of colleges. For example, Bound and Turner (2007) find that a 10 percent increase in a state's cohort size leads to a 4 percent decrease in the fraction of students earning a BA from that state. In the present context, if there are a fixed number of slots in the short run, the statewide effect of the policy should be weakly larger in the long run once supply can expand to meet demand and all new college-aspirants can attend.

It is also possible, however, given the DID design, that in the face of short-run capacity constraints, colleges could accept more applications from students in schools with no pre-policy center, displacing students enrolled at high schools with a pre-policy center. In this scenario, my estimated effect would reflect a short-run compositional effect, whereas the long-run DID estimate may be smaller as colleges expand and admit all students regardless of pre-policy test center status. Although I cannot conclusively rule out this story, there is little reason to think that in the matched sample of schools,

31. Results are the same for a more formal two-stage-least-squares analysis of the effect of taking the ACT on enrollment, where the excluded instrument is the interaction of a dummy for being in the post-policy period, with a dummy for being enrolled in a school without a pre-policy center.

students would be displaced at a higher rate from schools with a pre-policy test center than from schools without one. The two types of schools are similar across observed characteristics and have similarly sized supplies of college-goers pre-policy who could be potentially displaced by the new enrollees.

7. CONCLUSION

Nearly a dozen states have incorporated the ACT or SAT into their eleventh grade statewide assessment, requiring that all public school students take a college entrance exam. In this paper, I exploit the implementation of this policy to show that for every ten poor students who take a college entrance exam pre-policy and score college-ready, there are an additional five poor students who do not take the test but would score college-ready.

I compare changes in college-going rates pre- and post-policy among students at schools that did not have an ACT test center pre-policy to those that did, finding an increase in four-year enrollment by 0.6 percentage points or 2 percent. The effect was larger among boys (0.9 points), poor students (1.0 point), students in the poorest high schools (1.3 points), and students less likely to take a college entrance exam in the absence of the policy (1.3 points). The effect on enrolling in a four-year college for up to four years is similar, implying that students induced to attend college by the policy persist at the same rate as inframarginal college-goers.

Although these increases in the four-year college enrollment rate might not appear to be dramatically large, relative to other educational interventions this policy is inexpensive and currently being implemented on a large scale. The direct costs to states of a mandatory ACT policy include: (1) the per-student test fee, which for spring 2012 was \$32 (a \$2 discount off the price a student would pay privately);³² (2) a statewide administration management fee, which is approximately \$1 per student; and (3) the costs associated with trainings, meetings, and other logistical issues, which comes to less than \$1 per student.³³ Whereas (2) and (3) vary by state, the total cost is substantially less than \$50 per student in all mandatory ACT states, especially because the actual cost to a state is the direct cost of the policy minus the cost to design, administer, and grade the portions of the eleventh grade exam displaced by the ACT. Further, this cost calculation ignores savings to families who no longer have to pay for a college entrance exam. Thus, the “social cost” is even lower, given that much of the cost can be considered a transfer.

To show the relative cost-effectiveness of the mandatory ACT policy at increasing postsecondary attainment, I compare the policy to other educational interventions that increase college-going. I create an index of cost-effectiveness by dividing a policy’s cost by the proportion of students it induces into college. For example, assuming a \$50 per student cost and an increase in the four-year college enrollment rate of 0.6 percentage points, the amount spent by the mandatory ACT policy to induce a single child into

32. States can include the writing portion of the ACT for an additional \$15 per test.

33. All mandatory ACT costs come from communications between the author and staff at state departments of education. All costs of other policies are in 2007 dollars and come from Levine and Zimmerman (2010) unless otherwise noted. The costs of the early childhood programs and STAR have been discounted back to age zero using a 3 percent discount rate. Costs of mandatory ACT and other high school and college interventions have not been discounted.

college is \$8,333 ($= \$50 / 0.006$).³⁴ This figure is an upper bound, given that the true cost is substantially less than \$50 and the 0.6 percentage point effect is a likely lower bound. Also, targeting the policy at students in the poorest schools would reduce this figure to under \$4,000.

More traditional education policies are far more expensive than the mandatory ACT policy. Given the effects on college enrollment estimated in Deming (2009), Head Start has a cost per student induced into college of \$133,000 ($= \$8,000 / 0.06$). The cost per student induced into college from the class size decrease in the Tennessee STAR experiment is even larger: \$400,000 ($= \$12,000 / 0.03$) (Dynarski, Hyman, and Schanzenbach 2013). Dynarski (2003) showed that it takes approximately \$21,000 of traditional student aid to induce a single student into college, including the aid spent on students who would have enrolled regardless.

Other policies aim specifically to boost college enrollment by dismantling administrative barriers to enrollment. For example, Bettinger et al. (2012) randomly offered families at H&R Block assistance filling out the Free Application for Federal Student Aid, finding a cost per student induced into college of \$1,100 ($= \$88 / 0.08$). This policy is extremely cost effective, although it is unclear whether this policy could be successfully operated on a scale as large as the mandatory ACT policy.

Given that these estimated costs per student induced into college do not reflect the statistical precision of the enrollment effects, and that the interventions earlier in students' lives may have impacts beyond those on postsecondary attainment, these comparisons are best viewed as rough approximations. Nonetheless, they suggest that relative to other interventions operating on a large scale such as traditional student aid, the mandatory ACT policy is very cost effective.

Still, the mandatory ACT is far from a cure-all. The results in section 3 suggest that requiring all students to take a college entrance exam increases the supply of poor students scoring at a college-ready level by nearly 50 percent. Yet the policy increases the number of poor students enrolling at a four-year institution by only 6 percent. In spite of the policy, there remains a large supply of disadvantaged students who are high-achieving and not on the path to enrolling at a four-year college. Researchers and policy makers are still faced with the important question of which policies can further stem the tide of rising inequality in educational attainment.

ACKNOWLEDGMENTS

I thank Susan Dynarski, John Bound, Brian Jacob, and Jeff Smith for their advice and support. I am grateful for helpful conversations with Charlie Brown, Eric Brunner, Steve DesJardins, John DiNardo, Tom Downes, Rob Garlick, Michael Gideon, Andrew Goodman-Bacon, Steve Hemelt, Kevin Stange, Caroline Theoharides, Elias Walsh, and seminar participants at the University of Michigan and Association for Education Finance and Policy. Thanks for helpful comments from Amy Schwartz and two anonymous referees. I am grateful to ACT, Inc., and the College Board for the data used in this paper. In particular, I thank Ty Cruce, John Carrol, and Julie Noble at ACT, Inc., and Sherby Jean-Leger at the College Board. Thanks to the Institute of Education Sciences, U.S. Department of Education, for providing support through grant R305E100008 to the

34. One way to think of this calculation is as follows: if 1,000 students are treated with the policy at a cost of \$50 per student, six will be induced to attend college ($= 1,000 \times 0.006$) at a total cost of \$50,000 ($= \$50 \times 1,000$). Thus, the cost per student induced into college is \$8,333 ($= \$50,000 / 6$).

University of Michigan. Thanks to my partners at the Michigan Department of Education (MDE) and Michigan's Center for Educational Performance and Information (CEPI). This research used data structured and maintained by the Michigan Consortium for Education Research (MCER). MCER data are modified for analysis purposes using rules governed by MCER and are not identical to those data collected and maintained by MDE and CEPI. Results, information, opinions, and any errors are my own and are not endorsed by or reflect the views or positions of MDE or CEPI.

REFERENCES

- Abadie, Alberto, Matthew M. Chingos, and Martin R. West. 2012. Endogenous stratification in randomized experiments. NBER Working Paper No. 19742.
- ACT, Inc. 2002. *Understanding your ACT assessment scores*. Available www.act.org/content/act/en/products-and-services/the-act/your-scores/understanding-your-scores.html. Accessed 9 November 2016.
- Bailey, Martha J., and Susan M. Dynarski. 2011. Gains and gaps: A historical perspective on inequality in college entry and completion. In *Whither opportunity: Rising inequality, schools, and children's life chances*, edited by Greg Duncan and Richard Murnane, pp. 117–133. New York: The Russel Sage Foundation.
- Beshears, John, James J. Choi, David Laibson, and Bridgette C. Madrian. 2009. The importance of default options for retirement saving outcomes: Evidence from the United States. In *Social Security policy in a changing environment*, edited by Jeffrey Brown, Jeffrey Liebman, and David A. Wise, pp. 167–195. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226076508.003.0006.
- Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *Quarterly Journal of Economics* 127(3):1205–1242. doi:10.1093/qje/qjs017.
- Bound, John, and Sarah E. Turner. 2007. Cohort crowding: How resources affect collegiate attainment. *Journal of Public Economics* 91(5):877–899. doi:10.1016/j.jpubeco.2006.07.006.
- Bound, John, Michael Lovenheim, and Sarah E. Turner. 2010. Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. *American Economic Journal. Applied Economics* 2(3):1–31. doi:10.1257/app.2.3.129.
- Bowen, William G., Matthew M. Chingos, and Michael S. McPherson. 2009. *Crossing the finish line: Completing college at America's public universities*. Princeton, NJ: Princeton University Press.
- Bulman, George. 2015. The effect of access to college assessments on enrollment and attainment. *American Economic Journal. Applied Economics* 7(4):1–36. doi:10.1257/app.20140062.
- Busso, Matias, John DiNardo, and Justin McCrary. 2013. Finite sample properties of semiparametric estimators of average treatment effects. Unpublished paper, University of Michigan.
- Card, David. 1995. Earnings, schooling, and ability revisited. In *Research in labor economics*, vol. 14, edited by Solomon Polachek, pp. 23–48. Greenwich, CT: JAI Press.
- Carrell, Scott, and Bruce Sacerdote. (forthcoming). Why Do College Going Interventions Work? *American Economic Journal: Applied Economics*.
- Deming, David. 2009. Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal. Applied Economics* 1(3):111–134. doi:10.1257/app.1.3.111.

- Deming, David, and Susan M. Dynarski. 2010. Into college, out of poverty? Policies to increase the postsecondary attainment of the poor. In *Targeting investments in children: Fighting poverty when resources are limited*, edited by Philip Levine and David Zimmerman, pp. 283–302. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226475837.003.0011.
- Dillon, Eleanor, and Jeffrey Smith. 2017. The determinants of mismatch between students and colleges. *Journal of Labor Economics* 35(1):45–66.
- DiNardo, John. 2002. Propensity score reweighting and changes in wage distributions. Unpublished paper, University of Michigan.
- DiNardo, John, Nicole Fortin, and Thomas Lemieux. 1996. Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica* 64(5):1001–1044. doi:10.2307/2171954.
- Dynarski, Susan M. 2003. Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review* 93(1):279–288. doi:10.1257/000282803321455287.
- Dynarski, Susan M., Joshua M. Hyman, and Diane Whitmore Schanzenbach. 2013. Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management* 32(4):692–717. doi:10.1002/pam.21715.
- Dynarski, Susan M., Ken Frank, Brian Jacob, and Barbara Schneider. 2013. The effect of the Michigan Promise Scholarship on educational outcomes. Unpublished paper, University of Michigan.
- Dynarski, Susan M., Steven W. Hemelt, and Joshua M. Hyman. 2015. The missing manual: Using National Student Clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis* 37(1S):53S–79S. doi:10.3102/0162373715576078.
- Efron, Bradley, and Robert Tibshirani. 1993. *An introduction to the bootstrap: Monographs on statistics and applied probability*, vol. 57. New York: Chapman Hall. doi:10.1007/978-1-4899-4541-9.
- Eichler, Martin, and Michael Lechner. 2002. An evaluation of public employment programmes in the East German state of Sachsen-Anhalt. *Labour Economics* 9(2):143–186. doi:10.1016/S0927-5371(02)00039-8.
- Goodman, Sarena. 2016. Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics* 98(4):671–684. doi:10.1162/REST_a_00600.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies* 64(4):605–654. doi:10.2307/2971733.
- Hoxby, Caroline, and Christopher Avery. 2013. The “missing one-offs”: The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity* 46(1):1–65. doi:10.1353/eca.2013.0000.
- Hoxby, Caroline, and Sarah Turner. 2012. Expanding college opportunities for high-achieving, low income students. Stanford Institute for Economic Policy Research Discussion Paper No. 12–014.
- Hurwitz, Michael, Jonathan Smith, Sunny Niu, and Jessica Howell. 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. doi:10.3102/0162373714521866.

- Hyman, Joshua M. (forthcoming). Does money matter in the long run? Effects of school spending on educational attainment. [In press] *American Economic Journal: Economic Policy*.
- Imbens, Guido W., and Joshua Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62(2):467–475. doi:10.2307/2951620.
- Jackson, C. Kirabo. 2010. A little now for a lot later: A look at a Texas Advanced Placement Incentive Program. *Journal of Human Resources* 45(3):591–639.
- Klasik, Daniel. 2013. The ACT of enrollment: The college enrollment effects of state-required college entrance exam testing. *Educational Researcher* 42(3):151–160. doi:10.3102/0013189x12474065.
- Levine, Phillip B., and David J. Zimmerman. 2010. *Targeting investments in children: Fighting poverty when resources are limited*. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226475837.001.0001.
- Madrian, Brigitte C., and Dennis F. Shea. 2001. The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics* 116(4):1149–1187. doi:10.1162/003355301753265543.
- Pallais, Amanda. 2015. Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics* 33(2):493–520. doi:10.1086/678520.
- Pallais, Amanda, and Sarah Turner. 2006. Opportunities for low income students at top colleges and universities: Policy initiatives and the distribution of students. *National Tax Journal* 59(2):357–386. doi:10.17310/ntj.2006.2.08.