Effects of Test-Optional Admissions on Underrepresented Minority Enrollment and Graduation *

Trevor Osaki†

January 27, 2023

Abstract

A growing number of colleges and universities have made the submission of college entrance exam scores optional for undergraduate admissions to bolster racial diversity. This study uses a panel of liberal arts colleges from IPEDS and applies a two-way fixed effects approach to determine whether the policy is effective at achieving this goal. It also estimates the impact of the policy on the graduation rates for underrepresented minority (URM) students. It finds that the policy bolsters freshman URM enrollment among test-optional institutions throughout the sample as a whole, regardless of admissions selectivity and early versus late treatment timing. The effects of this policy on the URM 4-year and 6-year graduation rates are heterogeneous across colleges by their selectivity in admissions. While the most selective colleges in the panel experience no change in URM graduation rates, less-selective colleges experience declines in these graduation rates.

Keywords: test-optional policy, college and university admissions, college enrollment, underrepresented minority students, racial diversity, graduation rates

JEL: J08, I23, I24

---

*I would like to thank Kelly Bedard, Gonzalo Vazquez-Bare, Peter Kuhn, Sarah Robinson, Molly Schwartz, Chris Malloy, Woongchan Jeon, Ken Bao, and the Applied Economics Reading Group (ECON 290) for their invaluable suggestions and comments that are incorporated within this paper. All errors are my own.

†Department of Economics, University of California, Santa Barbara; tosaki@ucsb.edu.

1 Introduction

Racial and socioeconomic gaps in higher education attainment have been prevalent throughout the past several decades (Page and Scott-Clayton, 2016). The economic literature that addresses this issue includes studies that assess the impact of various policies intended to level the playing field, such as federal financial aid programs. One such policy, test-optional admissions, has received much less attention within this literature.

Many selective 4-year colleges and universities in the U.S. rely on SAT and ACT scores to assess students for undergraduate admission. However, a growing number of schools are making the submission of these test scores optional.[1] Although some of these schools reason that standardized test scores are unreliable indicators of college preparedness, many of them also suggest their new admissions policy can help promote racial diversity among their enrolled student bodies. Some test-optional schools cite anecdotal evidence of their policy's success in reaching this goal. For example, Providence College reports that its test-optional policy led to a 19 percent increase in underrepresented minority (URM) enrollees as well as a 56 percent increase in Pell-eligible enrollees (Epstein 2009).[2] As the number of test-optional schools grows, especially throughout the COVID-19 pandemic, the understanding of this policy's effects on diversity and graduation outcomes is becoming increasingly important.

This study uses a panel of private liberal arts colleges and applies two-way fixed effects to determine whether this test-optional policy successfully bolstered racial diversity at these institutions.[3] This study also examines the impact of the policy on the graduation rates for URM students, i.e., the share of URM students from a cohort who graduate within a specific duration of time. Liberal arts colleges are of interest since they primarily award 4-year undergraduate degrees (i.e., at least 50% of the degrees they offer to students), and a larger share of them dropped the test requirement within the panel's time frame compared to other types of institutions.[4]

[1] Specifically, the 2019 COVID-19 pandemic led the vast majority of colleges and universities to make the submission of standardized test scores optional due to the increased difficulty of facilitating proctored exams. Some of these schools effectively eliminated this requirement for admissions consideration. As discussed in Section 4, however, this study's scope does not encompass the pandemic.

[2] URM students refer to those from Black, Hispanic/Latino, or American Indigenous racial groups. These groups are noted to be underrepresented at selective colleges and universities.

[3] The test-optional policy is distinct from other similar policies, such as test-blind admissions (i.e., colleges do not consider test scores at all). Although the effects of this admission regime may be worth investigating, this study excludes test-blind institutions to isolate the effects of the test-optional policy. The data source used by this study contains information on only one such institution: Hampshire College.

[4] Liberal arts colleges compete with other types of 4-year institutions, e.g., private and public R1 and R2 universities, for undergraduate students seeking Baccalaureate degrees.

Proponents of the test-optional policy argue that standardized tests serve as admission barriers for prospective underrepresented students, resulting in racial and socioeconomic gaps in test scores. For example, *The Brookings Institute* reports that only 5% of Hispanic/Latino and 2% of Black students, compared to 60% Asian and 33% White students, are among the top-scoring SAT test-takers.[5] These gaps stem from a few channels. First, there are inequalities in access to resources that can help applicants within college admissions. Students from affluent backgrounds tend to have greater access to resources that can help them improve their scores on any standardized test via hiring a private tutor or registering for several administrations of the same exam (Vigdor and Clotfelter, 2003). Consequently, the SAT/ACT serves as a sorting mechanism that favors wealthy students and reinforces their disproportionate presence at the nation's most selective institutions (Anlon, 2009). These inequalities also lead to gaps in test scores across racial groups. Second, standardized college entrance exams are culturally biased (Freedle, 2003; Santelices and Wilson, 2010).[6]

There is some debate regarding the consequences of test-optional admissions. The most prominent point of contention toward this policy is the argument it could cause a decline in the academic preparedness of enrolling students (Epstein 2009). This corresponds to a phenomenon known as the *mismatch effect*, which posits that diversity-enhancing policies, such as race-blind admissions, can sort URM students into selective colleges and universities that would have otherwise rejected them. It hinges upon the assumption that some of these students tend to have lower academic credentials than their non-URM counterparts (Rothstein and Yoon, 2008). Consequently, upon enrollment, these students face peers that are much more academically prepared and will end up academically underperforming at their school of enrollment. Therefore, an analysis of the policy's effects on the graduation outcomes of URM students may, to some extent, shed light on whether the mismatch is a concern within a test-optional admissions regime.[7]

This study finds that test-optional admissions increase the volume of enrolling freshman students by 12.5%. Similarly, the policy increases the fraction of enrolling first-year students from a URM background by over 23% from the pre-treatment mean. These positive URM enrollment

[5]Richard V. Reeves and Dimitrios Halikias, *Race gaps in SAT scores highlight inequality and hinder upward mobility* (Washington D.C.: The Brookings Institute, 2017).

[6]Santelices and Wilson (2010) define *cultural bias* as the case in which cultural subgroups may interpret test items differently, especially in the case of *verbal* questions.

[7]The literature on mismatch has mixed results, with many of the utilized methodologies receiving heavy scrutiny. Arcidiacono et al. (2016) find robust evidence of school mismatch among STEM majors with the use of micro-data and counterfactual simulation. However, the appropriate methods needed to precisely identify mismatch exceed the scope of this study.

effects are evident across the entire sample, regardless of institutions' selectivity in admissions and whether they dropped the test requirement earlier versus later in the panel's time frame. It also finds that the policy has a negligible impact on the 4-year and 6-year URM graduation rates among highly selective colleges. On the other hand, relatively less-selective colleges experience a roughly 10% decline in these graduation rates. This result raises the possibility of college mismatch at these less-selective institutions.

A few studies use panel data sets of colleges & universities to estimate the impact of the policy on the overall share of students from a URM background on campus, and they find that it has little effect on this outcome (Belasco et al., 2015; Saboe and Terrizzi, 2019). Another recent study, Bennett (2022), uses a propensity-matched sample of colleges and universities from 2005 to 2015. It finds that the policy led to a minor increase in the share of URM students and the volume of applicants.

This study contributes to the literature in a few ways. First, it utilizes a sample incorporating a much larger share of test-optional schools (i.e., treated units). Thus, it backs out a relatively precise policy effect estimate. Second, unlike a few papers from the literature, this study estimates the policy's impact on the outcomes for *first-year* URM students (e.g., the fraction of first-year students from a URM background). These variables are likely informative since they are closely connected to the college admissions process and may be very responsive to the policy. Third, this study assesses the policy's impact on the graduation outcomes of URM students to determine whether students benefitting from the policy are more or less likely to persist beyond the first year of college. To the best of my knowledge, this is the first paper in the economic literature to assess this policy's impact on URM students' graduation outcomes. Finally, this study exploits heterogeneity in the policy's effects across colleges' degree of admissions selectivity.

This paper proceeds as follows. Section 2 discusses some background of the test-optional policy as well as the mechanisms that can explain how it could affect diversity outcomes of interest. Section 3 describes the sources of the data used in the primary analyses. Section 4 discusses the empirical strategy used to identify the effects of the policy. It also discusses the dynamic treatment effects of the policy as a test for identifying assumptions. Section 5 discusses the estimated effects of the test-optional policy. Finally, Section 6 concludes.

2 Background and Potential Mechanisms

The growing collection of schools that drop their standardized test requirement for admissions is informally known as the *test-optional movement*. This movement has primarily been driven by discourse that questions the usefulness of the SAT and ACT as proxies of student ability and cites their obstructive nature towards access to the most selective institutions. There indeed exists literature that substantiates these concerns about standardized tests. For example, Rothstein (2004) argues that the SAT's predictive power is lower than what the previous literature has suggested and recommends that it be assigned less importance within the admissions process. In addition, studies such as Blau, Moller, and Jones (2004), find that schools' reliance on standardized test scores can deter otherwise high-ability Black students from applying.

These concerns led Bowdoin Colleges and Bates College to drop their test score requirements in 1969 and 1984, respectively. Bates College, in particular, adopted it out of concern that its average SAT scores deterred strong students from applying (Epstein, 2009). It reported that this policy raised the number of applications it received and, at the same time, did not diminish the quality of its enrolling student cohorts. But it also found that the number of Black and Hispanic/Latino applicants increased, with almost half choosing not to submit their test scores. The success of early adopters, such as Bates College, led many other schools to drop their test requirements. Anecdotal evidence from test-optional colleges has caught the attention of other schools and has led them to adopt the policy. For example, the University of Chicago's vice president James Nondorf stated upon the university's recent adoption of its test-optional policy, "[the university's initiative] will further remove barriers to selective schools for students from underrepresented communities, including Pell applicants."[8] Other universities, such as Virginia Commonwealth University and George Mason University, have adopted the policy with the same intention.[9]

As shown in Figure J1, when a student navigates through the application for a test-optional college, they will be prompted to indicate whether they choose to submit their SAT or ACT score for admissions consideration. If a student opts to submit their test score, then it will be reviewed as a part of their application. Otherwise, if they do not, the college will heighten the importance of the remaining admissions criteria, such as high school GPA, letters of recommendation, or extracurricular activities. The fine print behind these colleges' admission processes is unobservable. However, the way each institution approaches test-optional admissions may vary. For example, as

[8] James G. Nondorf, "The University of Chicago, on Diversity," *The New York Times*, July 13, 2018.
[9] Joey Matthews, "VCU to Drop SAT Requirement," Richmond Free Press, January 1, 2015.

suggested by Figure J2, some colleges indicate that they provide equitable admission consideration to students regardless of whether they choose to submit their test scores or not. If an applicant withholds their test score, the college would infer their academic ability using remaining admissions criterion without penalizing them in any form. Other colleges indicate to applicants that with the absence of test scores, they will place greater weight on specific criteria when making an admission decision. For example, as shown in Figure J3, Denison University places additional weight on applicants' coursework rigor if they choose not to submit their test scores.

From an ex-ante perspective, the effects of the test-optional policy on the admission of URM students are ambiguous. As a first possibility, the policy could positively affect URM admission and enrollment. When colleges provide equitable admission consideration between submitters and non-submitters, they can curtail the issues of gaps in SAT and ACT scores. Thus, if many URM applicants are non-submitters, they can benefit from the policy and, therefore, face better probabilities of admission. Furthermore, the policy's positive impact on URM student enrollment could take the form of a *warming effect*. In this case, URM students value campuses that foster racial and ethnic diversity and gain the most utility from applying and enrolling at campuses in the presence of peers with similar backgrounds (Card and Krueger, 2005). Test-optional institutions may be signaling to prospective URM students that they are attempting to foster campus diversity by adopting the policy, so these students may be likely to apply and enroll there. Thus, this effect would bolster the volume and share of these students enrolling at these schools.

As a second possibility, the policy may have a negligible impact on URM admission and enrollment. As previously discussed, test-optional admissions may place greater weight on specific admission criteria, such as the availability of college preparatory coursework (e.g., AP/IB programs) or extracurricular activities. However, some school-specific criteria, such as the availability of college preparatory coursework and extracurricular activities, are noted to be unequally distributed across demographic groups (Espenshade & Radford, 2009; Iatarola, Conger, & Long, 2011; Klugman, 2013; Perna et al., 2013). Thus, these test-optional colleges could be replacing one inequitable set of criteria with another within their admissions processes. Consequently, not all URM applicants may be able to benefit from test-optional admissions.

As a final possibility, this policy could harm URM admission through mechanisms akin to signaling (Spence, 1973) and models of school admission (Avery & Levin, 2010; Pop-Echeles & Urquiola, 2013).[10] A college may desire to admit applicants with unobservable ability levels above

[10]These two studies consider the implications for applicants when they choose to reveal a piece of informa-

some cutoff, but it can only use noisy proxies such as grades and test scores to infer applicants' abilities. Under a test-optional regime, the college will only be able to use grades to infer a student's ability if they choose to withhold their test score. However, the strongest applicants with high test scores may submit them anyway to signal to the college that they meet the desirable ability level. But consequently, the college may wrongly infer students withholding test scores have some ability level below the cutoff, so they would face lower probabilities of admission. Thus, this policy would harm the admission and enrollment of URM students if they tend to withhold their test scores.[11]

Indeed, some of the mechanisms may be working in tandem with each other. Therefore, the sign of the estimated effects of the policy could roughly indicate which of these are predominant within the admissions and enrollment processes at test-optional schools.

3 Data and Sample

This study uses institution-specific panel data from the National Center of Education Statistics' Institutional Postsecondary Education Data System (IPEDS). The panel spans the academic years 2001-2002 through 2019-2020.[12] The data set includes the number of first-time URM students enrolling at each institution, the fraction of first-time students from a URM background, and the 4-year and 6-year URM graduation rates.[13] The data set also includes several control variables: logged full-time enrollment (FTE), logged tuition and fees, logged education and related (E & R) expenditures per FTE, logged institutional student grant aid per FTE, and a set of binary variables that reflects the extent college-preparatory classes are considered in admissions (e.g., college prep

tion (e.g., preference-based information) that may affect their placement relative to some admissions cutoff. For example, Avery and Levin (2010) consider early admissions within the context of college admissions. And Pop-Echeles & Urquiola (2013) consider students' choice of academic track within the Romanian secondary school system.

[11] This framework suggests other possible outcomes, or "equilibria." In particular, students with sufficiently high grades but low test scores could reveal them anyway to signal to college admissions that they possess the desired ability level. Hence, college admissions may infer that any student submitting their test scores is "high ability," whereas non-submitters would be inferred as "low ability."

[12] The choice of the initial period, 2001-2002, reflects data availability from IPEDS for some key variables. Although IPEDS includes periods beyond 2019-2020, the data set for this study does span further since the vast majority of schools dropped the test requirement during the 2020-2019 admission cycle due to the COVID-19 pandemic. This would essentially nullify the identification strategy discussed in Section 4 (i.e., almost all schools would be treated).

[13] *First-time students*, as defined by IPEDS, are students that have no prior postsecondary experience attending any institution for the first time at the undergraduate level. Therefore, none of these students transferred from a 2-year institution. Similarly, the outcomes on graduation solely reflect individuals that enrolled as first-time students.

classes are recommended or required).[14]

The complete list of test-optional institutions is obtained from *FairTest*, an organization that addresses fairness and accuracy issues within U.S. student test-taking. Although this list is convenient for identifying the institutions belonging to the test-optional (treatment) group, it does not provide the dates they adopted the policy. However, the exact adoption dates are obtained from IPEDS.[15] These dates are based on a yearly categorical variable that rates the extent to which test scores are considered in admissions.[16]

The sample includes a total of 149 liberal arts colleges, of which about 45% adopted the policy sometime between 2002-2003 and 2019-2020, respectively. Since test-optional institutions dropped the test requirement in various years, policy adoption is "staggered." Figure 1 contains the distribution for the years in which institutions from the sample adopted the policy. The inclusion of liberal arts colleges within the analyses ensures the comparability of all units across the treatment and control groups since a significant fraction of test-optional colleges through 2019-2020 belong to this category of institutions.[17] All of these institutions are either considered to be "selective," "more selective," or "most selective" by the U.S. News and World Report (USNWR). None of them are Historically Black Colleges and Universities (HBCUs).[18] Although IPEDS contains data on a larger number of institutions, many were discarded due to a substantial number of missing observations for some key variables. A large number of these excluded institutions were established after 2001-2002. Furthermore, all test-optional institutions (e.g., Bates College) that adopted their policy before 2001-2002 were dropped from the sample.[19] As a reference, Table A1 from Appendix

[14] All monetary control variables are logged and adjusted for inflation in 2019 USD. Furthermore, only a few colleges seem to have adjusted their consideration of college preparatory courses simultaneous with dropping the test-requirement. However, the sample size is relatively small and, therefore, sensitive to outliers. Thus, these binary variables are included in the specification, although their inclusion changes the point estimates by a very small amount.

[15] The set of schools suggested by IPEDS to be test-optional is consistent with that of FairTest.

[16] This variable takes on a value of "1" if test scores are required, "2" if they are recommended, "3" if they are neither required nor recommended, or "5" if they are considered but not required. Any institution whose variable takes on a value of 2, 3, or 5 is considered test-optional.

[17] There are other types of institutions that adopted test-optional admissions, such as doctoral universities. However, these institutions differ from liberal arts colleges in many ways, such as endowment, enrollment level, etc. Therefore, these institutions may not serve as proper control units with the analyses. Furthermore, most of these types of institutions dropped the test requirement within the last several periods leading up to 2019-2020. On the other hand, the timing of treatment among liberal arts colleges has more variation within the time frame of the panel.

[18] A preliminary sample of institutions from IPEDS contains three HBCUs, none of which are test-optional. Appendix G reproduces much of the analyses using a sample that includes these institutions. It shows that the inclusion of HBCUs do not significantly change the results from the body of this paper.

[19] IPEDS has data on a total of seven colleges thought would be considered as "always-treated" units. As discussed in the next section, this study estimates the effects of the policy using a TWFE approach with

A contains the list of test-optional institutions in this sample and their year of policy adoption.

The summary statistics for the outcome and control variables are displayed in Table 1. These statistics correspond to the 2001-2002 (pre-treatment) observations of the displayed variables. Columns (1) and (2) contain the statistics for the test-optional and test-requiring institutions, respectively. Column (3) includes p-values for the difference in means between the test-optional and requiring institutions for all variables. These differences, except for a few control variables, are statistically indistinguishable from zero.

4 Empirical Strategy

4.1 Specifications

This study utilizes a two-way fixed effects (TWFE) approach. The institution's choice to adopt the test-optional policy serves as the "treatment." As discussed in Section 3, treatment is staggered across academic years. The coefficients are estimated via Ordinary Least Squares (OLS). Given institution i in academic year $t \in \{2001\text{-}2002, ..., 2019\text{-}2020\}$, the primary specification is represented by the following equation:

$$y_{it} = \beta P_{it} + \mathbf{X}'_{it}\gamma + \alpha_i + \lambda_t + \epsilon_{it} \tag{1}$$

where y_{it} is some outcome variable of interest (e.g., the logged number of first-time URM students). Standard errors are clustered by institution.

Specifically, P_{it} takes on a value of "1" in any academic year when an institution's matriculating class is affected by the policy. For example, if an institution adopted the policy during the 2015-2016 academic year for the incoming class of 2020, it is first indicated as test-optional in the 2016-2017 academic year. When y_{it} denotes graduation outcomes, P_{it} is lagged by 6 periods to appropriately reflect how cohorts of students are affect by the policy.[20]

The term α_i represents institutional fixed effects while λ_t represents academic year fixed effects.

staggered treatment. The estimator for these effects is a weighted average of average treatment effects on the treated (ATT) (Goodman-Bacon, 2020). When always-treated units are included within such regression, they are treated as control units and are given disproportionately greater weight. If their treatment effects change over time, they bias TWFE estimator from a meaningful parameter.

[20] Graduation data on IPEDS reflects cohorts enrolling 6 years prior to data reporting. For example, graduation rate data from 2016-2017 reflects the cohort entering in 2010-2011. So, P_{it} is also lagged by 6 periods when the outcome variable is the 4-year graduation rate.

Institutional fixed effects should control for time-invariant factors that may confound the estimates of the policy's effects. The vector $\boldsymbol{X}_{it}$ includes all of the control variables. Some of these variables, such as institutional grants per FTE, can control for some additional policy changes that may accompany the adoption of the test-optional policy, such as financial aid expansion, and therefore confound its effects.[21]

Section 5.2 explores the potential issue of heterogeneity in the policy's effects across colleges' admissions in selectivity. This subsection splits test-optional institutions into two groups: "selective" and "highly selective" institutions. To perform this analysis, this study estimates a regression model similar to equation (1):

$$y_{it} = \beta P_{it} + \delta(S_i \cdot P_{it}) + \boldsymbol{X}'_{it}\gamma + \alpha_i + \lambda_t + \epsilon_{it}. \tag{2}$$

It includes an interaction term $S_i \cdot P_{it}$ where S_i takes on a value of "1" if institution i is highly selective in admissions. Thus, $\beta + \delta$ represents the effects of the policy for these types of institutions while β alone captures the the effects for selective institutions. δ captures the difference in effects between these two types of institutions. The vector $\boldsymbol{X}$ includes the same covariates as in equation (1) plus those saturated in S_i. α_i and λ_t are the same fixed effects from equation (1), although the latter is also saturated in S_i.

Finally, Section 5.3 compares the policy's effects on first-time URM enrollment outcomes between colleges dropping the test-optional policy early in the panel versus later. To distinguish between early and later adopters of this policy, this study estimates the following specification that is also similar to equation (1):

$$y_{it} = \beta_1(P_{it} \cdot E_i) + \beta_2(P_{it} \cdot L_i) + \boldsymbol{X}'_{it}\boldsymbol{\gamma} + \alpha_i + \lambda_t + \epsilon_{it}. \tag{3}$$

Here, y_{it} corresponds to either the logged number of first-time URM students or the fraction of first-time students of a URM background. E_i is a binary variable that takes on a value of "1" if test-optional institution i adopted the policy early within the time frame. Similarly, L_i takes on a value of "1" if a test-optional institution adopted the policy later. Both E_i and L_i always equal "0" for all test-requiring institutions. Hence, the test of $\beta_1 = \beta_2$ can indicate whether the policy's

[21] As previously discussed, the adoption of the test-optional policy has been noted to be paired with financial aid expansion. To my knowledge, there are no other known types of policy changes that have accompanied the adoption of this policy.

effects between early and late-adopters are distinguishable.

4.2 Dynamic Treatment Effects & Identifying Assumptions

This study estimates the dynamic effects of the test-optional policy on the outcome variables of interest to check for pre-treatment trends and the common trends assumption. Specifically, this study estimates the following regression specification:

$$y_{it} = \sum_{\tau=-7}^{-2} \delta_\tau D_{it} + \sum_{\tau=0}^{7} \mu_\tau D_{it} + \boldsymbol{X}'_{it}\gamma + \alpha_i + \lambda_t + \epsilon_{it}. \tag{4}$$

Similar to equation (1), y_{it} represents an outcome variable, $\mathbf{X}_{it}$ represents the same vector of controls, and α_i and λ_t represent institution and year fixed effects, respectively. δ_τ and μ_τ represent the leading and lagging coefficients across event time τ, i.e., the relative number of periods to adoption of the test-optional policy. This equation omits $\tau = -1$, so the leads and lags represent the dynamic effects of the policy relative to one period prior to adoption. The leading and lagging periods are symmetric across event time.[22] However, since the adoption of the test-optional policy is staggered, the test-optional institutions from the sample are unbalanced in these periods.

5 Results

5.1 Average effects among all colleges

Figures 2-5 contain the plotted dynamic effect estimates across event time for all outcome variables. Table 2 contains the p-values from the joint tests on the leading coefficients across each outcome (i.e., these test $\delta_\tau = 0$ for $-7 \leq \tau \leq -2$). These tests fail to find evidence that at least one leading coefficient is statistically different from 0. That said, the common trends assumption is presumed to hold for all outcome variables.

Recent literature suggests the coefficients for the leads and lags may be contaminated by effects from other periods when the treatment is staggered (Sun and Abraham, 2020). Consequently, this issue may invalidate the joint test on the leads. To address the above issue, this study re-estimates the dynamic treatment effects using alternative estimators proposed by de Chaisemartin

[22] The chosen number of periods is roughly based on the median number of leading and lagging periods across treated units from the sample.

and D'Haultfœuille (2021) that are robust to treatment heterogeneity. Further discussion on these dynamic effects can be found in Appendix D, which shows that the results of an analogous test also presume that the common trends assumption holds.

Columns (1) and (2) of Table 3 contain the point estimates of the TWFE coefficient from equation (1) for outcomes on first-time URM enrollment. It indicates that the policy raises the fraction of first-time students from a URM background by about 0.0188 points (i.e., on average, test-optional colleges experience a 23.5% increase from the pre-treatment average). Also, these estimates show that the number of first-time URM students rises by 12.5%. This suggests that the increased fraction can be primarily attributed to an increased inflow of first-time URM students enrolling at these institutions rather than a drop in the overall volume of enrollees.[23]

Although the average effects of the policy on URM enrollment are positive, the effects for each test-optional college in the sample may vary by size. This study investigates further by calculating the difference in first-time URM enrollment outcomes between two periods after and before policy adoption for all test-optional colleges within the sample. Figure 6 displays these differences for the share of first-time students from a URM background within a histogram. It illustrates positive differences among the majority of test-optional colleges in the sample. Figure 7 also indicates the abundance of positive differences for the logged number of first-time URM students enrolling at these colleges. These figures collectively suggest that most test-optional colleges are likely to implement equitable consideration between submitters and non-submitters of test scores and that they experience warming effects from URM students.

All in all, these results suggest that the policy is, to a small extent, effective at bolstering racial diversity at liberal arts colleges. As a possibility, however, these patterns may be driven by an overall increase in the volume of enrolling students, regardless of race. In that case, the increase in the shares of URM students may be a byproduct of that goal. However, Appendix F discusses the effects of the policy on the logged number of first-time students regardless of race, and it suggests that this is not likely to be the case.

Finally, columns (3) and (4) of Table 3 displays the test-optional policy's effects on the graduation outcomes of URM students. The point estimate for each outcome is relatively small (e.g., a 0.02 point decrease in the 4-year URM graduation rate) and statistically insignificant. These results suggests that the policy has a negligible impact on the 4-year and 6-year graduation rates

[23]As an exercise, Appendix E shows that the policy has little impact on the enrollment outcomes of non-first-time URM students.

for URM students. So, on average, racial diversity improvements resulting from the policy may not diminish beyond the first year of college. However, further discussion of these effects across school selectivity is provided in the next subsection.

The TWFE approach to estimating these effects have a few caveats. First, colleges and universities may voluntarily elect to drop their test-requirement. Thus, the “treatment” takes the form of an endogenous policy change. This could be an issue if colleges with relatively lower shares of URM students *self-select* themselves into this treatment. However, the treated and control colleges from the sample have comparable and statistically indistinguishable shares of URM students on their campuses (i.e., the difference is 0.004 with $p = .62$).[24] Also, as shown in Appendix B, the estimates are robust to the exclusion of control variables. Finally, this study re-estimates equation (1) using a propensity-trimmed sample and finds that the resulting point estimates are comparable to that of the full sample. The logistic regression used to construct this sample suggest that the pre-treatment share of URM students on campus is not a significant predictor for adopting the test-optional policy. Some further discussion behind this propensity trimmed sample can be found in Appendix C.

Alternatively, the endogeneity of policy adoption may stem from colleges within each state following neighboring institutions in dropping the test requirement. This study collapses the panel to the state level and conducts analogous state-level analysis to investigate whether this phenomenon is evident among colleges in the sample. Specifically, it estimates the impact of the share of students attending a test-optional college within each state on state-level freshman URM enrollment outcomes. The results of this exercise can be found in Appendix I. They indicate that the estimates of the share of treated students on state-level URM enrollment outcomes are noisy. In short, there is little evidence that colleges’ decision to drop the test requirement reflects the actions of neighboring in-state institutions.

Second, recent papers show that the TWFE estimator with staggered treatment is a weighted sum of all possible 2×2 difference-in-differences estimators (Goodman-Bacon, 2021). However, some of these weights may be negative in the presence of treatment heterogeneity which could render the estimator uninterpretable (de Chaisemartin and D’Haultfœuille, 2020). The TWFE estimators for the URM enrollment outcomes (e.g., the fraction of first-time students that are URM) are composed of 536 ATTs, of which 5.6% of them receive negative weights. These weights

[24]This share corresponds to the overall fraction of students on campus that are from a URM background, regardless of freshman status.

sum up to about -0.012. On the other hand, the TWFE estimators associated with the graduation rates contain zero negative weights. Appendix D discusses estimates for the URM enrollment outcomes using an alternative estimator that is robust to treatment heterogeneity, and it suggests that negative weighting is not a significant issue.

5.2 Differences in Admissions Selectivity

The analyses, thus far, encompass institutions regardless of their degree of selectivity in admissions. Highly selective institutions tend to place a relatively higher weight on test scores within their admissions processes (Marin and Horn, 2008). Therefore, as a possibility, the effects of the test-optional policy may differ at colleges considered to be more selective than others. So, the effects of the policy may be more salient at the most selective colleges. To investigate this possibility, this study distinguishes institutions that are "selective" and "highly selective" (i.e., are considered more selective and most selective by the USNWR). The sample contains 54 selective institutions and 95 highly selective institutions of which 44.4% and 45.3% are test-optional, respectively. The summary statistics for these two sub-samples can be found in Tables A2 and A3.[25]

Equation (2) is estimated to distinguish the effects of the policy across institutional selectivity. The point estimates are displayed in Table 4 across all outcome variables. Interestingly, among the URM enrollment outcomes, the point estimates for selective colleges among the URM enrollment outcomes are higher than that for highly selective colleges. However, the differences between these two sub-groups are insignificant. Appendix F shows that neither of these sub-groups is likely to be adopting the test-optional policy to increase overall enrollment, regardless of students' racial background. Although the effect of the policy on the first-time enrollment volume at selective colleges is significant, it is much smaller than the effect on first-time URM enrollment.

The point estimates of the 4-year and 6-year URM graduation rates are negative and significant among selective colleges (e.g., the 4-year graduation rate drops by 0.0678 points, or 12.8% from the pre-treatment level). On the other hand, the point estimates for highly selective colleges are small and insignificant, and in fact, the differences between the two sub-samples are significant. In other words, the policy led to a decline in the URM graduation rates at selective colleges. At these

[25]The mean URM graduation rates for selective colleges are larger than that of highly selective colleges, with the difference for 6-year graduation rate being statistically significant. However, this difference within the propensity-trimmed sample from Appendix C is slightly smaller but statistically insignificant. The results of the analyses in that section is very comparable to the ones discussed below (i.e., under the full sample).

institutions, students benefitting from the policy may be less likely to graduate.[26] Consequently, the graduation rates for URM students decrease as a result of the policy.

Further discussion on these graduation rate effects is provided in Section 6. Appendix H shows that test-optional admissions have little impact on graduation rates for non-URM students (e.g., White and Asian students) at both selective and highly selective colleges.

5.3 Timing of Policy Adoption

As discussed, test-optional institutions within the sample adopted the policy throughout different years between 2002-2003 and 2019-2020. However, the effects of the policy on URM enrollment may vary across these institutions by the timing of treatment, i.e., whether they dropped the test requirement relatively early within the panel's time frame or later. As one possibility, the warming effects on enrollment may be more prevalent among colleges adopting the policy relatively early in the panel. Specifically, if colleges drop the test requirement early within an environment where there are few or no other test-optional colleges, they would have an easier time attracting prospective URM students seeking test-optional admissions. Consequently, colleges that drop the test requirement relatively later in this time frame may experience little policy effects on URM enrollment because prospective students seeking test-optional admissions already have an alternative: the early adopting colleges. Thus, the policy's impact on racial diversity at late-adopting colleges may be negligible if they are crowded out by early adopters.

This study distinguishes institutions that adopted the policy *early* in the panel versus *later* by estimating equation (3) with first-time URM enrollment outcome variables. It considers early adopters as colleges that dropped the test requirement anytime up to 2010, i.e., the midpoint of the time frame. All other test-optional colleges that dropped their test requirement after 2010 are considered to be late adopters.

Table 5 contains the point equations of equation (3), where each column corresponds to an outcome variable. The point estimates show that the average effects on first-time URM enrollment outcomes between institutions adopting the policy during the first versus the second half of the time frame are comparable and statistically indistinguishable. All in all, these results indicate that there is no strong evidence that the policy's effects differ between early and late adopters.[27]

[26] This could be backed by the fact that the mean pre-treatment admission rate of selective colleges is higher than that of highly selective colleges by about 22%.

[27] This study considers other ways to define early adopters, such as those that drop the test requirement through 2007 (i.e., the first 25% of adopters). However, under this alternative threshold, the effects from

Appendix B reproduces the analysis of Table 5 without the use of control variables, and it illustrates similar patterns. Appendix F shows that neither early (i.e., adopting through 2007) nor late adopters experienced a significant increase in enrolling first-time students, regardless of race.

6 Discussion

This study finds that the test-optional policy effectively bolsters racial diversity at liberal arts colleges. It also finds that the policy has little impact on the URM graduation rates at highly selective colleges. But it finds evidence that the policy led to a decline in graduation rates at relatively less-selective colleges. Finally, this study finds that the URM enrollment effects among colleges dropping the test requirement early in the panel versus later are comparable and indistinguishable.

Given the sign of the point estimates, these results are consistent with a warming effect. While most residential liberal arts colleges today are not necessarily highly selective or elitist, individuals are inclined to equate a residential liberal arts education with an "elite" form of higher education (Astin, 1999). Therefore, these institutions tend to be associated with the perception of prestige/selectivity. Thus, their adoption of this policy may send a signal to prospective URM students that they are trying to facilitate a welcoming environment for them.

These results suggest that the mismatch effect may be present at selective colleges versus highly selective ones. Indeed, the most selective institutions tend to draw higher shares of applicants from the top quintile of their graduating high school cohort (Bound et. al, 2009). Thus, high-ability students who presumably face greater probabilities of retention and successful graduation may sort themselves to these institutions, even within a test-optional admissions regime. However, the relationship between graduation rates and the test-optional policy is unlikely to be causal. These estimates may encompass other unobservable policies and phenomena that affect retention, such as the availability of advising and outreach at highly selective colleges or the lack thereof at relatively less-selective institutions for students benefitting from the policy.[28] Therefore, there is no certainty whether the test-optional policy alone led to the decline in the URM graduation rates at selective colleges and whether mismatch is a significant issue within test-optional admissions.

early and late adopters are also statistically indistinguishable.

[28]The estimates of the policy's effects on URM enrollment may not be exempt from this issue as well. As discussed in Footnote 17, however, the test-optional policy is likely paired with financial aid expansion or shifts in weighting toward college preparatory courses, which would have been captured by the control variables.

As suggested by the results, these test-optional colleges may have implemented an enhanced, equitable admissions process for submitters and non-submitters of test scores. However, the outcome of interests used in this analysis reflects enrollment. Although enrollment effects can pose implications for admissions, it would do so, albeit to a limited extent. Unfortunately, IPEDS does not disaggregate its admissions data by demographics. Thus, richer data is necessary to fully understand how the policy affects the admissions processes for URM students at these institutions.

Furthermore, as discussed in Section 2, some test-optional colleges place greater weight on inequitable admissions criteria, such as high school rigor. Therefore, not all URM applicants may be benefitting from the policy. However, the estimated effects of the policy on URM enrollment outcomes suggest that this channel is dominated by warming effects or the prevalence of equitable admissions within the sample.

Certainly, the policy's effects on liberal arts colleges could yield implications on its beneficiaries' post-graduation returns. A recent study, Carnevale et. al (2020), indicates that the median long-run (40-year) net present value (NPV) returns of liberal arts colleges is approximately \$918,000, i.e., almost \$200,000 higher than the median for all types of 4-year colleges and universities.[29] Also, it finds that the returns to the most selective liberal arts colleges (i.e., most of the institutions within the "highly-selective" category of the sample) are higher at \$1,135,000. Since the policy does not seem to diminish the graduation rates of URM students at these institutions, students benefitting from the policy could experience significant returns on investment and possibly upward social mobility.[30] On the other hand, the negative impact of the policy on the URM graduation rates at relatively less-selective institutions may be problematic. Carnevale et. al (2020) also shows that high graduation rates are correlated with high investment returns.[31] Therefore, students who benefit from the test-optional policy but face lower probabilities of graduation attainment will not be able to realize these gains.

Some questions about the test-optional policy remain for future investigation. By its design,

[29]Other types of colleges & universities, as designated by the Carnegie Classification, include doctoral-granting universities and special-focus schools.

[30]Carnevale et. al (2020), contains the estimated 40-year returns for a significant amount of individual liberal arts colleges. By using these figures, this current study finds comparable trends within its sample. The median 40-year return to investment among all 149 institutions is about \$949,000 which also exceeds the average for all types of institutions. However, the median is higher at \$999,000 compared to \$882,500 for selective institutions.

[31]This finding is also consistent with patterns observed within this current study's sample. The correlation between the estimated 40-year return to investment of colleges included in the sample and the 2020 graduation rates is 0.691.

the policy's effects should be salient to students from financially disadvantaged backgrounds. Although IPEDS data on such students is limited, an analysis of this policy on *socioeconomic* diversity in admission, enrollment, and graduation would be informative in understanding its effects.[32] Next, other types of institutions, such as doctoral-granting universities, are increasingly dropping their test requirement, so future studies may be able to assess the impact of the policy on racial or even socioeconomic diversity at these types of schools. The analyses of public institutions may especially be of interest since many of these systematically target a large and relatively diverse array of students, so the impact of the policy on these schools may be pronounced.[33] Finally, future analyses can shed light on how the test-optional policies shift enrollment patterns between different types of institutions (e.g., private versus public or selective versus less-selective institutions).[34]

Racial and economic gaps in college enrollment and attainment continue to persist, despite the relatively high returns to post-secondary education and significant student aid efforts. This study primarily contributes to the discussion of how policies can bolster racial diversity within colleges and universities and considers the test-optional policy as a measure that can achieve this goal. Interestingly, many colleges and universities that dropped the test requirement during the COVID-19 pandemic are either prolonging their admission regimes or making them permanent.[35] As the costs and benefits of this policy come to light, colleges and universities will become better informed on whether to maintain their admissions regime. With the ongoing prevalence of test-optional admissions among colleges and universities, knowledge of these costs and benefits is becoming more relevant than ever.

[32]Although IPEDS contains data on the number of first-year Pell grant recipients at each college, the availability of that variable is constrained to a small number of periods.

[33]For example, the 1960 California Master Plan for Education stipulates that the University of California (UC) system, which elected to permanently drop its test requirement in 2020, targets the top 12.5% of California high school students. In fact, within the past few years, the UC system has been accommodating roughly 18% of California high school students, and they consider expanding further to match the growing demand for admission. (Source: https://www.latimes.com/california/story/2021-07-27/california-is-failing-to-meet-demand-for-uc-admission-why-its-a-crisis)

[34]For example, Hinrichs (2012) and Backes (2012) are studies that assess the impact of an admission policy (i.e., affirmative action bans) on the enrollment patterns of URM students. They find that the policy shifted the fraction of URM students attending more selective colleges to less-selective ones. They posit that the ban also led to an increase in enrollment of these students at public 2-year colleges, but they find little empirical evidence of that occurring, mostly owed to data limitations.

[35]For example, Occidental College, which dropped its test requirement at the onset of the pandemic, notes on its website that it has no plans to return to test-requiring admissions in the future.

References

Alon, S., "The evolution of class inequality in higher education: Competition, exclusion, and adaption," *American Sociological Review*, 2009, *74* (5), 731–755.

Arcidiacono, Peter, Esteban M. Aucejo, and Joseph V. Hotz, "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California," *American Economic Review*, 2016, *106* (3), 525 – 562.

_ , Michael Lovenheim, and Maria Zhu, "Affirmative Action in Undergraduate Education," *Annual Review of Economics*, 2015, *7*, 487 – 518.

Astin, Alexander W., "How Liberal Arts Colleges Affect Students," *Daedalus*, 1999, *128* (1), 77 – 100.

Avery, Christopher and Jonathan Levin, "Early Admissions at Selective Colleges," *American Economic Review*, 2010, *100* (5), 2125 – 2156.

Backes, Ben, "Do Affirmative Action Bans Lower Minority College Enrollment and Attainment? Evidence from Statewide Bans," *Journal of Human Resources*, 2012, *47* (2), 435 – 455.

Belasco, Andrew S., Kelly O. Rosinger, and James C. Hearn, "The Test Optional Movement at America's Selective Liberal Arts Colleges: A Boon for Equity or Something Else?," *Education Evaluation and Policy Analysis*, 2015, *37* (2), 206 – 223.

Bennett, C.T., "Untested admissions: Examining changes in application behaviors and student demographics under test-optional policies," *American Educational Research Journal*, 2022, *59*, 180–216.

Blau, J.R., S. Moller, and LV Jones, "Why test? Talent loss and enrollment loss," *Social Science Research*, 2004, *33* (3), 409 – 434.

Bound, John, Brad Hershbein, and Bridget Terry Long, "Playing the admissions game: Student reactions to increasing college competition," *Journal of Economics Perspectives*, 2009, *23* (4), 119 – 146.

Card, David and Alan B. Krueger, "Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas," *ILR Review*, 2005, *58* (3), 416–434.

Carnevale, Anthony P., Ban Cheah, and Martin Van Der Werf, "ROI of Liberal Arts Colleges: Value Adds Up Over Time," *Georgetown University Center on Education and the Workforce*, 2020, (14).

de Chaisemartin, Clément and Xavier D'Haultfoeuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economics Review*, 2020, *110* (9), 2964—-2996.

_ and _ , "Difference-in-Differences Estimators of Intertemporal Treatment Effects," 2021. Working paper.

Epstein, Jonathan, "Behind the SAT-Optional Movement: Context and Controversy," *Journal of College Admission*, 2009, *204*, 8–19.

Freedle, R.O., "Correcting the SAT's ethnic and social-class bias: A method for re-estimating SAT Scores," *Harvard Education Review*, 2003, *73*, 1 – 43.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225*, 254—-277.

Hinrichs, Peter, "The Effects of Affirmative Action Bans on College Enrollment, Education Attainment, and the Demographic Compositions of Universities," *Review of Economics and Statistics*, 2012, *94* (3), 712–722.

Jacobsen, Louis, Robert Lalonde, and Danielle Sullivan, "Earnings Losses of Displaced Workers," *American Economic Review*, 1993, *83* (4), 685–709.

Marin, Patricia and Catherine L. Horn, "Realizing Bakke's Legacy: Affirmative Action, Equal Opportunity, and Access to Higher Education," 2008.

McPherson, Michael S. and Morton Owen Shapiro, "The Future Economic Challenges for the Liberal Arts Colleges," *Daedalus*, 1999, *128* (1), 47 – 75.

Page, Lindsay C. and Judith Scott-Clayton, "Improving college access in the United States: Barriers and policy responses," *Economics of Education Review*, 2015, *51*, 493–520.

Pallais, Amanda, "Small Differences That Matter: Mistakes in Applying to College," *Journal of Labor Economics*, 2016, *33* (2), 4–22.

Pop-Echeles, Cristian and Miguel Urquiola, "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 2013, *103* (4), 1289 – 1324.

Rothstein, Jesse, "College performance predictions and the SAT," *Journal of Econometrics*, 2004, *121* (1-2), 297–317.

— **and Albert H Yoon**, "Affirmative Action in Law School Admissions: What Do Racial Preferences Do?," *The University of Chicago Law Review*, 2008, *75*, 649–714.

Saboe, M. and S. Terrizzi, "SAT optional policies: Do they influence graduate quality, selectivity or diversity?," *Economic Letters*, 2019, *174*, 13–17.

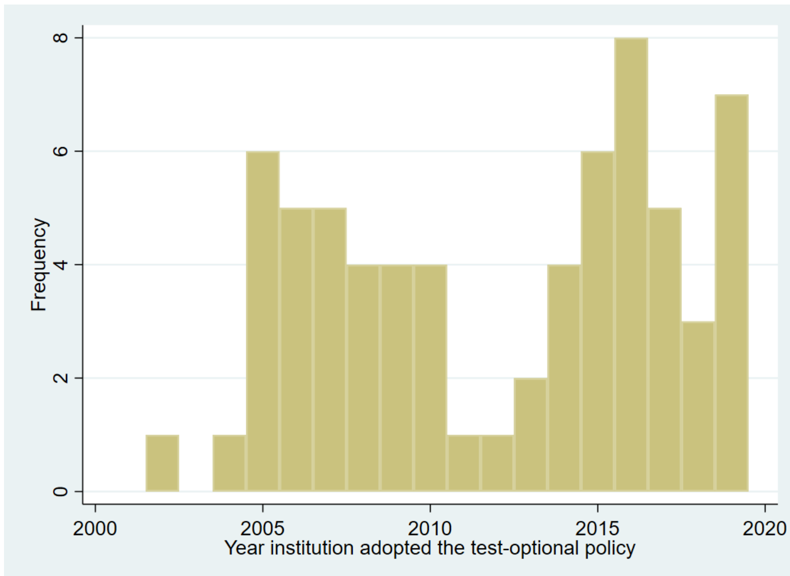
Santelices, M.V. and Mark Wilson, "Unfair Treatment? The Case of Freedle, the SAT, and the Standardization Approach to Differential Item Functioning," *Harvard Educational Review*, 2010, *80* (1), 106 – 134.

Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics*, 1974, *87* (3), 281 – 306.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, *225*, 175–199.

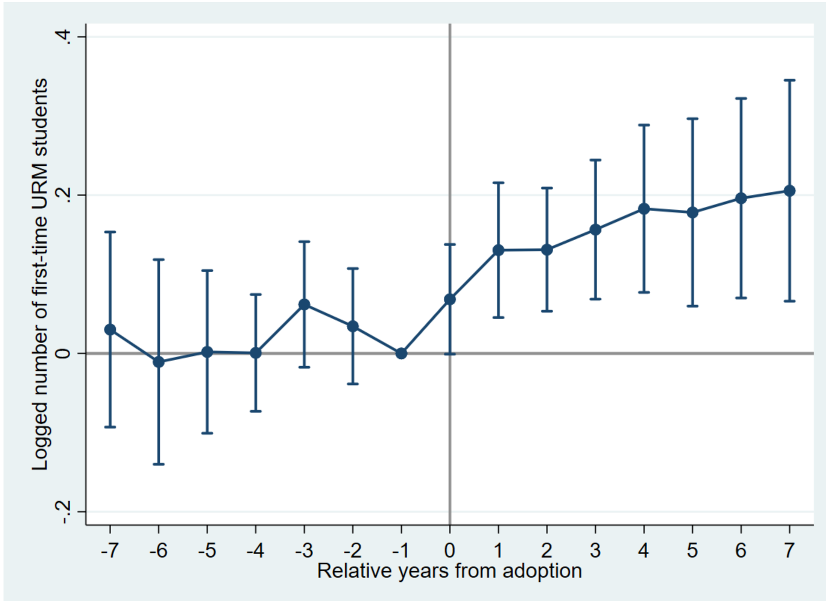
Figures

Figure 1: Distribution of Policy Adoption Year



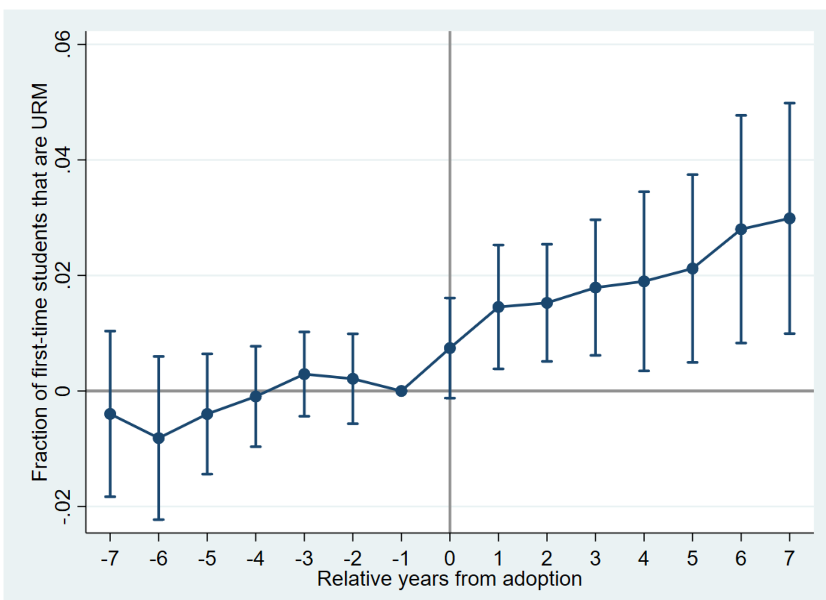
Notes: This figure displays the distribution of the years in which test-optional institutions from the sample adopted their admissions policy.

Figure 2



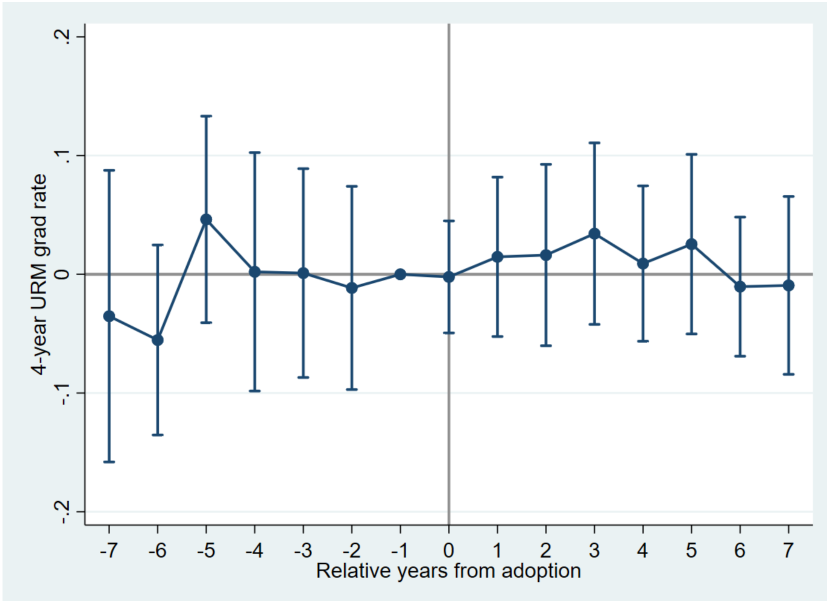
Notes: This figure illustrate the dynamic effects on the logged number of first-time URM students. Coefficient estimates from equation (4). are plotted across event time (i.e., years relative to the test-optional policy taking effect). They are represented by the blue dots. The accompanying bands represent the 95% confidence intervals of these coefficients. Figures 2-6 are arranged similarly, but they reflect different outcome variables.

Figure 3



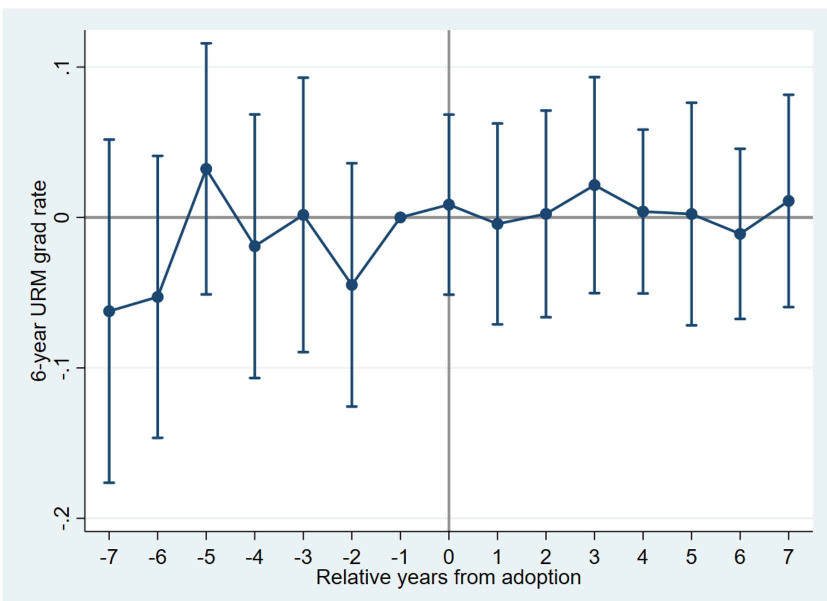
Notes: This figure illustrates the estimated dynamic effects of the test-optional policy on the fraction of first-time students enrolling at liberal arts colleges that are of URM status.

Figure 4



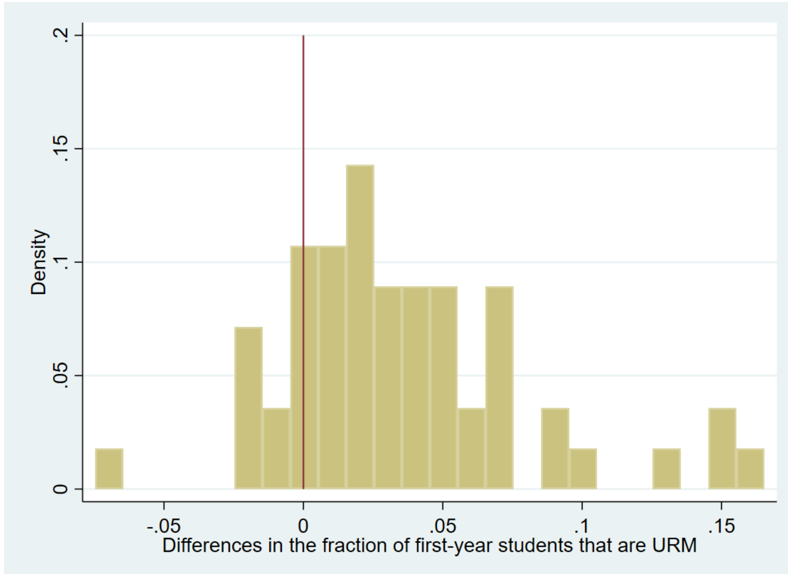
Notes: This figure illustrates the estimated dynamic effects of the test-optional policy on the 4-year graduation rate for URM students.

Figure 5



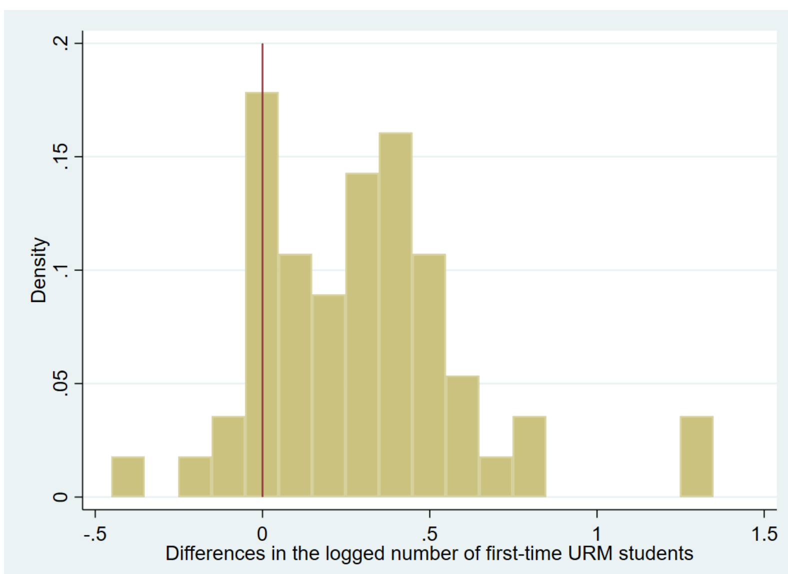
Notes: This figure illustrates estimated dynamic effects of the test-optional policy on the 6-year graduation rate for URM students.

Figure 6



Notes: For each test-optional college in the sample, this study calculates the difference in the share of first-time students from a URM background two periods after and before policy adoption (i.e., $Y_i^{2+} - Y_i^{2-}$). This figure is a histogram for these differences, where the vertical red line corresponds to "0" difference.

Figure 7



Notes: For each test-optional college in the sample, this study also calculates the difference in the logged number of first-time URM students background two periods after and before policy adoption (i.e., $Y_i^{2+} - Y_i^{2-}$). This figure is a histogram for these differences, where the vertical red line corresponds to "0" difference.

Tables

Table 1: Summary Statistics

	Test-optional (1)	Test-requiring (2)	p-value of diff. (3)
First-time URM students	31.00	29.87	0.77
	(20.39)	(25.90)	
Fraction of first-time students that are URM	0.08	0.08	0.74
	(0.05)	(0.05)	
4-year URM grad rate	0.53	0.52	0.78
	(0.19)	(0.24)	
6-year URM grad rate	0.63	0.60	0.48
	(0.16)	(0.23)	
Tuition & Fees	30,290	27,957	0.03
	(5,547)	(7,729)	
Full-time enrollment	1,489	1,372	0.28
	(597.9)	(714.9)	
Institutional grants per FTE	9.70	20.00	0.03
	(4.73)	(43.03)	
E & R expenditures per FTE	28.41	69.92	0.03
	(17.99)	(166.56)	
College prep courses not considered	0.03	0.02	0.84
	(0.17)	(0.16)	
College prep courses recommended	0.36	0.62	0.00
	(0.48)	(0.49)	
College prep courses required	0.60	0.34	0.00
	(0.49)	(0.48)	
Observations	67	82	–

Notes: Columns (1) and (2) contain summary statistics for test-optional and test-requiring institutions using the 2001-2002 (i.e., pre-treatment) observations. Standard deviations are in parenthesis. Column (3) contains the p-values from the difference means between these two groups. These p-values are clustered by institution. None of the variables are logged. Therefore, the means for tuition & fees, E & R expenditures per FTE, and institutional grants per FTE are in terms of 2019 dollars. The sample size varies slightly across each variable due to non-reporting.

Table 2: Results of Joint Test of Leads

	p-value for joint F-test
a) Logged number of first-time URM Students	0.588
b) Fraction of first-time students that are URM	0.823
c) 4-year URM graduation rate	0.385
d) 6-year URM graduation rate	0.276

Notes: This table displays the p-values for the joint test in leading coefficients from equation (4) across all outcome variables of interest. The point estimates of the leads and lags are displayed in Figures 2-6.

Table 3: Primary Results

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	4-year URM graduation rate (4)
Test-optional	0.125**	0.0188**	-2.32e-05	0.00450
	(0.0419)	(0.00628)	(0.0142)	(0.0131)
Observations	2,820	2,824	2,783	2,812

Notes: This table displays the results from estimating equation (1) on all outcome variables, and it solely reports the TWFE estimate of the test-optional policy. Columns (3) and (4) correspond to graduation outcomes, so the regressions for those use a lagged treatment indicator (i.e., $P_{i,t-6}$). One star indicates a 5% significance level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution.Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

Table 4: Effects of the Policy across Selectivity

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	6-year URM graduation rate (4)
Test-optional	0.154*	0.0247*	-0.0678**	-0.0476**
	(0.0690)	(0.0121)	(0.0153)	(0.0149)
Highly Selective × Test-optional	-0.0422	-0.00792	0.0820**	0.0662**
	(0.0831)	(0.0139)	(0.0211)	(0.0200)
Observations	2,827	2,831	2,790	2,819

Notes: This table displays the point estimate from equation (2), which distinguishes institutions by their degree of admission selectivity. The row labeled with "Test-optional" contains the point estimate for the coefficient of the treatment indicator of having the policy in place. The row labeled "Highly selective × Test-optional" contains the point estimate for the coefficient of the interaction term between the treatment indicator and another indicator for being highly selective in admissions. Each column corresponds to an outcome variable. Similar to Table 1, the point estimates for the coefficients of the control variables are not reported here. One star indicates a 5% level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

Table 5: Effects of the Policy by Adoption Timing

	Logged Number of freshman URM students (1)	Fraction of freshman students that are URM (2)
Early Adopter	0.117	0.0208*
	(0.0597)	(0.0102)
Late Adopter	0.133**	0.0164*
	(0.0516)	(0.00814)
p-value	0.830	0.751
Observations	2,820	2,824

Notes: This table displays the point estimate from equation (3), which distinguishes the effects of the policy by adoption timing. The row labeled "Early adopter" corresponds to the estimated effect of the policy on institutions that dropped the test requirement early. Similarly, the row labeled "Later adopter" corresponds to the effect on institutions adopting the policy later. Each column corresponds to a first-time URM outcome variable. The p-values on the bottom row reflect the test of $\beta_1 = \beta_2$. One star indicates a 5% level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

A Sample Details

Table A1: List of Test-Optional Liberal Arts Colleges

Year	School	State	Year	School	State
2002	Mount Holyoke College	MA	2014	Trinity College	CT
2004	Pitzer College	CA	2014	Beloit College	WI
2005	St. Lawrence University	WI	2014	Wesleyan University	CT
2005	Knox College	IL	2015	University of Puget Sound	WA
2005	Juniata College	PA	2015	The College of Idaho	ID
2005	Lawrence University	NY	2015	Kalamazoo College	OH
2005	College of the Holy Cross	MA	2015	Emanuel College	MA
2006	Susquehanna University	PA	2015	Transylvania University	PA
2006	Franklin & Marshall College	PA	2015	Allegheny College	PA
2006	Drew College	NJ	2016	Whittier College	CA
2006	Bennington College	VT	2016	Skidmore College	NY
2006	Gustavus Adolphus College	MN	2016	Warren Wilson College	NC
2007	Lake Forest College	IL	2016	Whitman College	WA
2007	Gettysburg College	PA	2016	Willamette University	OR
2007	Wittenberg University	OH	2016	Houghton University	NY
2007	Hobart William Smith College	NY	2016	Cornell College	IA
2007	Denison University	OH	2016	Presbyterian College	SC
2008	Stonehill College	MA	2017	Hanover College	IN
2008	Guilford College	NC	2017	Linfield University	OR
2008	Goucher College	MD	2017	Wells College	NY
2008	Augustana College	IL	2017	Ripon College	WI
2009	Albright College	PA	2017	Bloomfield College	NJ
2009	Agnes Scott College	GA	2017	Wofford College	SC
2009	Smith College	MA	2018	Doane University	NE
2009	Ursinus College	PA	2018	Birmingham-Southern College	AL
2010	The University of the South	TN	2018	Austin College	TX
2010	Washington & Jefferson College	PA	2019	Spring Hill College	AL
2010	Lycoming College	PA	2019	Bucknell University	PA
2010	Saint Michael's College	VT	2019	Southwestern University	TX
2011	Saint Anselm College	NH	2019	Randolph College	VA
2012	Earlham College	IN	2019	Hendrix College	AK
2013	Ohio Wesleyan University	OH	2019	Monmouth College	NJ
2013	Washington College	MD	2019	DePauw University	IN
2014	Bryn Mawr College	PA			

Notes: This table provides a list of included in the sample. The "Year" column correspond to each institution's year of adopting their test-optional policy. For example, Pitzer College's policy was effective in the Fall 2004 round of admissions.

Table A2: Summary Statistics (selective colleges)

	Test-optional (1)	Test-requiring (2)	p-value of diff. (3)
First-time URM students	26.54	20.20	0.27
	(20.10)	(21.81)	
Fraction of first-time students that are URM	0.08	0.06	0.26
	(0.06)	(0.05)	
4-year URM grad rate	0.45	0.34	0.04
	(0.17)	(0.21)	
6-year URM grad rate	0.57	0.42	0.01
	(0.18)	(0.21)	
Tuition & Fees	26,617	22,681	0.00
	(4,492)	(5,300)	
Full-time enrollment	1,267	1,088	0.18
	(469)	(494)	
Institutional grants per FTE	10.16	19.47	0.16
	(5.29)	(35.11)	
E & R expenditures per FTE	29.25	49.47	0.16
	(22.80)	(73.85)	
College prep courses not considered	0.04	0.00	0.32
	(0.20)	(0.00)	
College prep courses recommended	0.29	0.63	0.01
	(0.46)	(0.49)	
College prep courses required	0.67	0.37	0.03
	(0.48)	(0.49)	
Observations	24	30	–

Notes: This table is equivalent to Table 1, except it solely reflects colleges within the sample that are considered to be “selective” in admissions. The sample size varies slightly across each variable due to non-reporting.

Table A3: Summary Statistics (highly selective colleges)

	Test-optional (1)	Test-requiring (2)	p-value of diff. (3)
First-time URM students	33.49	35.44	0.69
	(20.36)	(26.61)	
Fraction of first-time students that are URM	0.08	0.09	0.58
	(0.05)	(0.05)	
4-year URM grad rate	0.57	0.62	0.24
	(0.19)	(0.19)	
6-year URM grad rate	0.66	0.71	0.14
	(0.14)	(0.17)	
Tuition & Fees	32,340	31,000	0.29
	(5,028)	(7,287)	
Full-time enrollment	1,613	1,536	0.60
	(631)	(773)	
Institutional grants per FTE	9.45	20.31	0.10
	(4.43)	(47.32)	
E & R expenditures per FTE	27.95	81.72	0.06
	(14.95)	(201.43)	
College prep courses not considered	0.02	0.04	0.67
	(0.15)	(0.19)	
College prep courses recommended	0.40	0.62	0.03
	(0.49)	(0.49)	
College prep courses required	0.56	0.33	0.02
	(0.50)	(0.47)	
Observations	43	52	–

Notes: This table is equivalent to Table 1, except it solely reflects colleges within the sample that are considered to be "highly selective" in admissions. The sample size varies slightly across each variable due to non-reporting.

B Results without Control Variables

This section reproduces the estimates for Tables 3-5, but it excludes the use of control variables while retaining the institution and academic year fixed effects. The results for these exercises are provided in Tables B1-B3. In short, the point estimates are similar to the ones in Tables 3-5. Thus, the results are not sensitive to the exclusion of control variables (i.e., these variables merely improve the precision of the TWFE estimates).

Table B1: Primary Results – without control variables

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	4-year URM graduation rate (4)
Test-optional	0.123**	0.0191**	-0.00272	0.00580
	(0.0406)	(0.00644)	(0.0125)	(0.0119)
Observations	2,827	2,831	2,790	2,819

Notes: This table displays the results from estimating equation (1) on all outcome variables, but with a specification excluding control variables. The row labeled with "Test-optional" contains the point estimate of TWFE coefficient. Each column corresponds to an outcome variable. Columns (3) and (4) correspond to graduation outcomes, so the regressions for those use a lagged treatment indicator (i.e., $P_{i,t-6}$). One star indicates a 5% significance level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

Table B2: Effects of the Policy across Selectivity – without control variables

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	6-year URM graduation rate (4)
Test-optional	0.154*	0.0247*	-0.0678**	-0.0476**
	(0.0690)	(0.0121)	(0.0153)	(0.0149)
Highly Selective × Test-optional	-0.0422	-0.00792	0.0820**	0.0662**
	(0.0831)	(0.0139)	(0.0211)	(0.0200)
Observations	2,827	2,831	2,790	2,819

Notes: This table displays the point estimate from equation (2) without control variables. The row labeled with "Test-optional" contains the point estimate for the coefficient of the treatment indicator of having the policy in place. The row labeled "Highly selective × Test-optional" contains the point estimate for the coefficient of the interaction term between the treatment indicator and another indicator for being highly selective in admissions. Each column corresponds to an outcome variable. The point estimates for the coefficients of the control variables are not reported in this table. One stars indicate a 5% level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

Table B3: Effects of the Policy by Adoption Timing – without control variables

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)
Early adopter	0.121*	0.0215*
	(0.0586)	(0.0106)
Late adopter	0.127*	0.0179*
	(0.0510)	(0.00791)
p-value	.939	.799
Observations	2,827	2,831

Notes: This table displays the point estimate from equation (3), but with a specification that excludes control variables. The row labeled "Early adopter" corresponds to the estimated effect of the policy on institutions that dropped the test requirement early. Similarly, the row labeled "Later adopter" corresponds to the effect on institutions adopting the policy later. Each column corresponds to a first-time URM outcome variable. The p-values on the bottom row reflect the test of $\beta_1 = \beta_2$. One star indicates a 5% level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

C Propensity-Trimmed Sample

This section replicates the estimates of equations (1) and (2) using a sample derived from a propensity trimming technique suggested by Crump et al. (2009).[36] A propensity-trimmed sample is constructed in a few steps. First, this study estimates a logit model of the probability that an institution adopted the policy as a function of the 2001-2002 (pre-treatment) observations of the covariates used in equation (1), as well as the fraction of students that are from a URM background (i.e., regardless of freshman status), and the 6-year URM graduation rate. Motivated by Section 5.2, this logit model is estimated for selective and highly selective institutions. Second, it calculated the fitted values of these regressions, which serve as each institution's estimated propensity for adopting the test-optional policy. Third, it discards all institutions whose propensity score lies outside of [0.1,0.9]. Therefore, this method effectively excludes institutions that have virtually no probability of adopting as well as those that will almost certainly do so.

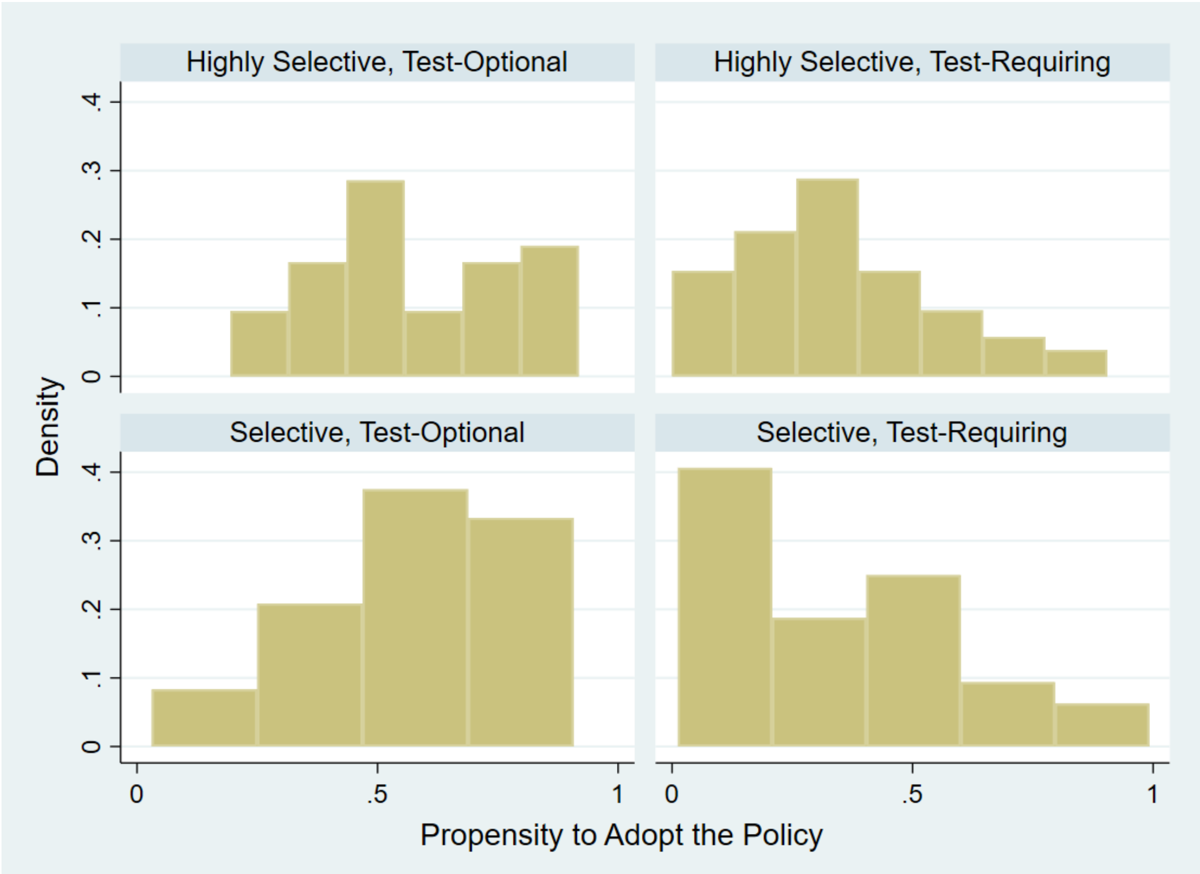
Figure C1 contains the densities of the institutions' propensity scores, which were calculated by the logit regressions, across treatment groups and selectivity levels. Furthermore, the results of these regressions can be found in Table C1.[37] Interestingly, they show that the coefficients for overall campus diversity and the 6-year URM graduation rate are insignificant. Therefore, these variables may not be strong predictors of adopting the test-optional policy, regardless of selectivity.

The resulting propensity-trimmed sample contains 130 institutions, of which 44 and 86 are selective and highly selective, respectively (i.e., 19 institutions are dropped due to having propensities less than 0.10 or greater than 0.90). In each of these sub-samples, at least 48% of institutions are test-optional. The summary statistics for the full sample, selective institutions, and highly selective institutions can be found in Tables C2, C4, and C5. Compared to Table 4, the differences in mean for the 4-year and 6-year URM graduation rates among selective colleges are relatively balanced. Although the differences are sizable (e.g., approximately 0.08 for the 6-year URM graduation rate), they are statistically indistinguishable from zero. Finally, Tables C3 and C6 replicate Tables 3 and 4 using the propensity-trimmed sample. The point estimates between these sets of tables are comparable in magnitude. Thus, the estimates from Tables 3 and 4 are not biased by the inclusion of institutions that were either likely to adopt the test-optional policy or were unlikely to do so.

[36]This technique was originally proposed to address a lack of overlap in the distribution of outcome and control variables. This issue could lead to estimates being substantially biased and having large variances.

[37]None of the selective institutions required college prep courses in 2001, the logit regression does not estimate a coefficient for the corresponding indicator variable.

Figure C1



Notes: This figure displays the distribution of institutions' propensity to adopt the test-optional policy across selectivity and treatment group (i.e., test-optional versus test-requiring schools).

Table C1: Logit Regressions

	Selective (1)	Highly selective (2)
Fraction of students that are URM	7.629	6.097
	(5.042)	(7.195)
6-year graduation rate	3.329	-1.565
	(2.099)	(2.003)
Logged tuition & fees	7.962	1.455
	(4.943)	(1.886)
Logged full time enrollment	-2.347	1.042
	(1.816)	(0.597)
Logged institutional grants per FTE	-1.352	2.768*
	(1.221)	(1.202)
Logged E & R expenditures per FTE	-1.993	-2.329*
	(1.927)	(1.028)
College prep courses recommended	-0.211	-0.940
	(0.775)	(1.629)
College prep courses required		0.136
		(1.635)
Observations	54	95

Notes: This table contains the results of the logistic regressions used to construct the propensity trimmed sample. These regressions incorporate 2001-2002 (pre-treatment) observations. They were run separately for selective and highly selective colleges. The point estimates for the former and latter are in columns (1) and (2), respectively. One star corresponds to a 5% significance level while two stars correspond to a 1% level. Standard errors are in parenthesis and clustered by institution.

Table C2: Summary Statistics

	Test-optional (1)	Test-requiring (2)	p-value of diff. (3)
First-time URM students	31.19	30.53	0.87
	(20.77)	(23.23)	
Fraction of first-time students that are URM	0.08	0.08	0.68
	(0.05)	(0.05)	
4-year URM grad rate	0.53	0.55	0.64
	(0.19)	(0.20)	
6-year URM grad rate	0.63	0.63	0.98
	(0.16)	(0.20)	
Tuition & Fees	30,351.33	29,355.47	0.32
	(5,396.33)	(6,036.80)	
Full-time enrollment	1,494.63	1,475.33	0.86
	(606.94)	(677.58)	
Institutional grants per FTE	9.71	18.53	0.11
	(4.80)	(43.90)	
E & R expenditures per FTE	28.49	62.10	0.11
	(18.34)	(167.21)	
College prep not considered	0.03	0.03	0.98
	(0.18)	(0.17)	
College prep recommended	0.36	0.58	0.01
	(0.48)	(0.50)	
College prep required	0.59	0.38	0.01
	(0.50)	(0.49)	
Observations	64	66	–

Notes: Columns (1) and (2) contain summary statistics for test-optional and test-requiring institutions using the 2001-2002 (i.e., pre-treatment) observations. However, this table incorporates the propensity-trimmed sample. Standard deviations are in parenthesis. Column (3) contains the p-values from the difference means between these two groups. These p-values are clustered by institution. None of the variables are logged. Therefore, the means for tuition & fees, E & R expenditures per FTE, and institutional grants per FTE are in terms of dollars. The sample size varies slightly across each variable due to non-reporting.

Table C3: Primary Results under Propensity-Trimmed Sample

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	4-year URM graduation rate (4)
Test-optional	0.124**	0.0169**	0.00773	0.00843
	(0.0413)	(0.00634)	(0.0145)	(0.0135)
Observations	2,459	2,463	2,428	2,454

Notes: This table displays the results from estimating equation (1) on all outcome variables. However, these estimates are based on the propensity-trimmed sample The row labeled with "Test-optional" contains the point estimate of the TWFE coefficient. Each column correspond to an outcome variable. Columns (3) and (4) correspond to graduation outcomes, so the regressions for those use a lagged treatment indicator (i.e., $P_{i,t-6}$). One star indicates a 5% level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

Table C4: Summary Statistics (selective colleges)

	Test-optional (1)	Test-requiring (2)	p-value of diff. (3)
First-time URM students	26.95	24.09	0.68
	(20.90)	(24.38)	
Fraction of first-time students that are URM	0.08	0.07	0.52
	(0.06)	(0.06)	
4-year URM grad rate	0.45	0.42	0.59
	(0.18)	(0.17)	
6-year URM grad rate	0.58	0.50	0.16
	(0.18)	(0.17)	
Tuition & Fees	26,584.54	24,644.84	0.11
	(3,764.73)	(4,025.50)	
Full-time enrollment	1,279.45	1,219.14	0.67
	(488.19)	(448.53)	
Institutional grants per FTE	10.19	16.80	0.36
	(5.42)	(33.01)	
E & R expenditures per FTE	29.00	40.03	0.47
	(23.83)	(67.37)	
College prep not considered	0.05	0.00	0.32
	(0.21)	(0.00)	
College prep recommended	0.27	0.55	0.07
	(0.46)	(0.51)	
College prep required	0.68	0.45	0.13
	(0.48)	(0.51)	
Observations	22	22	–

Notes: This table is equivalent to Table A2, but it reflects the "selective" colleges from the propensity-trimmed sample. The sample size varies slightly across each variable due to non-reporting.

Table C5: Summary Statistics (highly selective colleges)

	Test-optional (1)	Test-requiring (2)	p-value of diff. (3)
First-time URM students	33.40	33.75	0.94
	(20.60)	(22.21)	
Fraction of first-time students that are URM	0.08	0.08	0.98
	(0.05)	(0.05)	
4-year URM grad rate	0.58	0.61	0.36
	(0.19)	(0.18)	
6-year URM grad rate	0.66	0.70	0.26
	(0.14)	(0.17)	
Tuition & Fees	32,324.42	31,710.78	0.59
	(5,087.73)	(5,490.42)	
Full-time enrollment	1,607.33	1,603.43	0.98
	(637.27)	(738.50)	
Institutional grants per FTE	9.45	19.40	0.18
	(4.49)	(48.77)	
E & R expenditures per FTE	28.22	73.13	0.14
	(15.02)	(199.18)	
College prep not considered	0.02	0.05	0.59
	(0.15)	(0.21)	
College prep recommended	0.40	0.59	0.09
	(0.50)	(0.50)	
College prep required	0.55	0.34	0.05
	(0.50)	(0.48)	
Observations	42	44	–

Notes: This table is equivalent to Table B3, but it reflects the highly "selective colleges" from the propensity-trimmed sample. The sample size varies slightly across each variable due to non-reporting.

Table C6: Effects of the Policy across Selectivity (under Propensity-Trimmed Sample)

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	6-year URM graduation rate (4)
Test-optional	0.160*	0.0243*	-0.0543*	-0.0543*
	(0.0723)	(0.0117)	(0.0229)	(0.0216)
Highly selective × Test-optional	-0.0230	-0.00578	0.0707*	0.0730**
	(0.0857)	(0.0138)	(0.0272)	(0.0253)
Observations	2,573	2,577	2,542	2,568

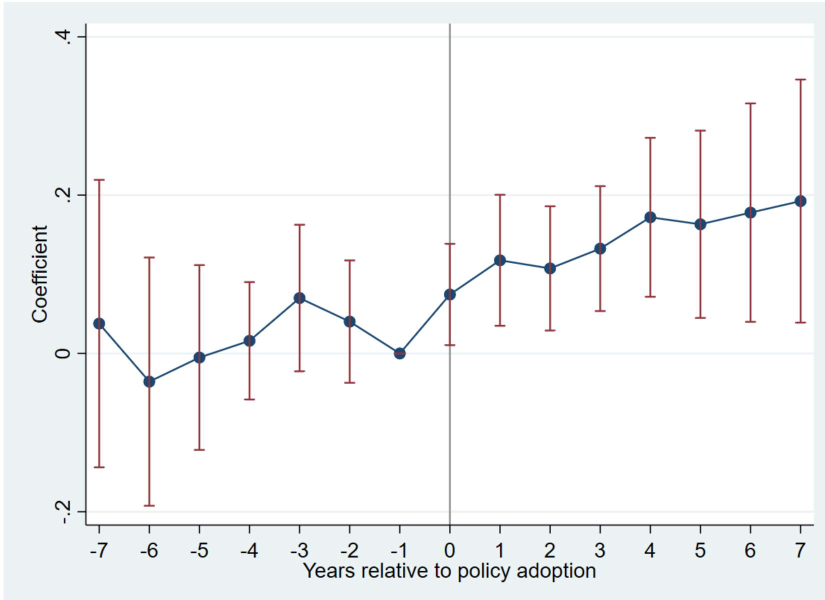
Notes: This table displays the point estimate from equation (2), which distinguishes institutions by their degree of admission selectivity. However, these correspond to the propensity-trimmed sample. The row labeled with "Test-optional" contains the point estimate for the coefficient of the treatment indicator of having the policy in place. The row labeled "Highly selective × Test-optional" contains the point estimate for the coefficient of the interaction term between the treatment indicator and another indicator for being highly selective in admissions. Each column corresponds to an outcome variable. The point estimates for the coefficients of the control variables are not reported in this table. One star indicates a 5% significance level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

D Heterogeneity Robust Estimation

This section replicates Figures 2-6 using alternative estimators suggested by de Chaisemartin and D'Haultfœuille (2021). These estimators, denoted as DID_ℓ, are robust to treatment heterogeneity and dynamic effects. So, in place of leads and lags, this study estimates dynamic and placebo effects, respectively. Therefore, a joint placebo test serves as an analogous and robust test for the common trends assumption. Figures D1-D4 contain the plotted placebo and dynamic effect estimates across "event time" (i.e., the period relative to adoption) for all outcome variables. The p-values for the joint placebo tests are displayed in Table D1 across each outcome variable. Similar to the joint F-test of the leads, these tests also fail to find strong evidence that at least one placebo is statistically different from 0. Thus, the common trends assumption is presumed to hold under this specification.

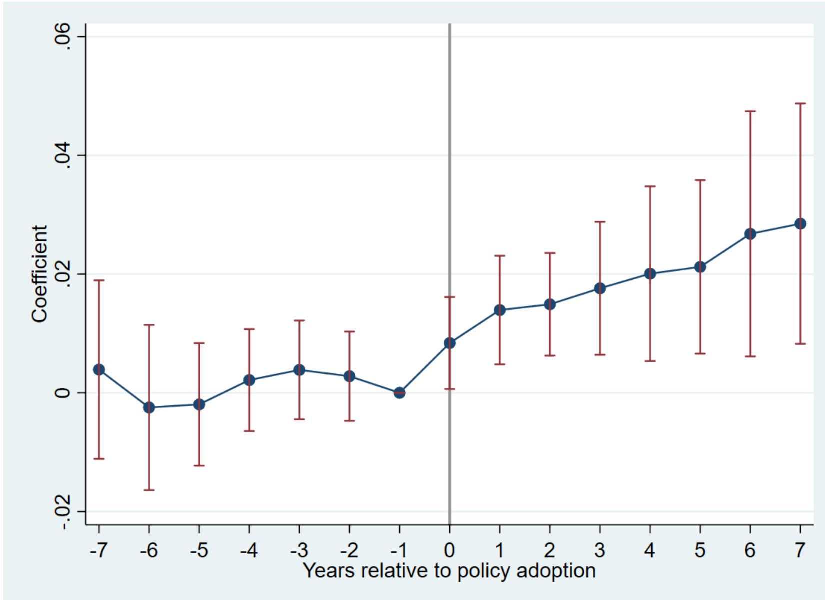
This section also re-estimate the URM enrollment effects from columns (1)-(4) of Table 3 using the average of the dynamic DID_ℓ estimates discussed above. These serve as analogues to the original two-way fixed effects estimates from Table 3. The average effect for each outcome is reported in Table D2, with bootstrapped standard errors displayed in parenthesis. The point estimates are mostly comparable to the ones from the primary two-way fixed effects specification, so, therefore, heterogeneous treatment effects may not be a significant issue. Although the DID_ℓ point estimates for the URM graduation rates differ from that of TWFE, they are imprecise.

Figure D1



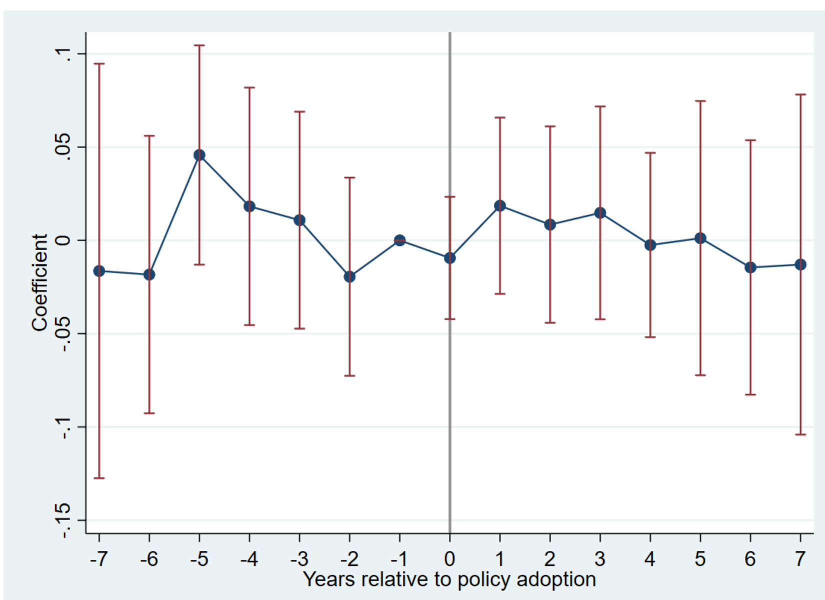
Notes: This figure illustrate the computed placebo and dynamic effects of the test-optional policy on the logged number of URM students enrolling for the first-time at liberal arts colleges. The placebo and dynamic effects are analogous to the leads and lags of an event-study. These estimates are plotted across "event time" (i.e., the number years relative to the test-optional policy being adopted). They are represented by the dots. The accompanying bands represent the 95% confidence intervals of these estimates, which were bootstrapped across 100 replications. All proceeding figures are arranged similarly, but they reflect different outcome variables.

Figure D2



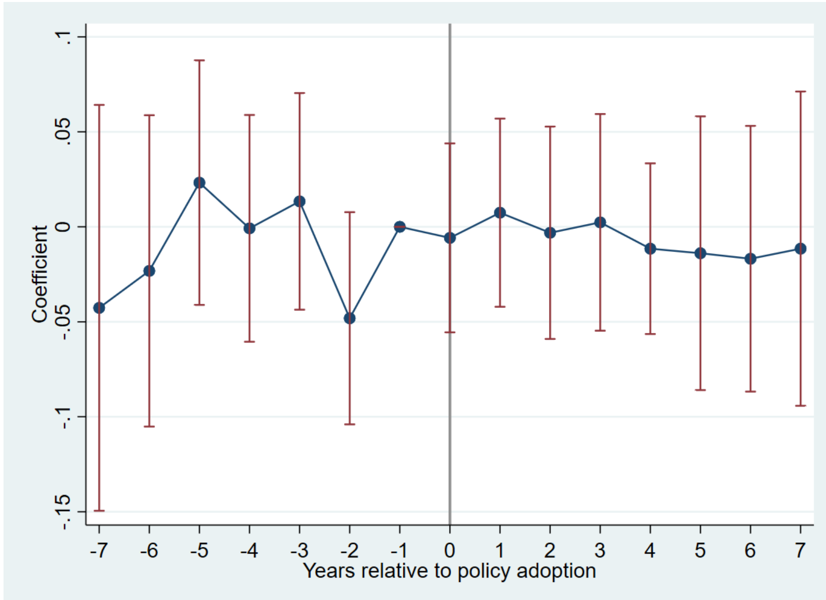
Notes: This figure illustrates the computed placebo and dynamic effects of the test-optional policy on the fraction of first-time students enrolling at liberal arts colleges that are of URM status.

Figure D3



Notes: This figure illustrates the computed placebo and dynamic effects of the test-optional policy on the 4-year graduation rate for URM students.

Figure D4



Notes: This figure illustrates the computed placebo and dynamic effects of the test-optional policy on the 6-year graduation rate for URM students.

Table D1: Results of Joint Placebo Test

	p-value for joint placebo test
a) Logged number of first-time URM Students	0.561
b) Fraction of first-time students that are URM	0.913
c) 4-year URM graduation rate	0.212
d) 6-year URM graduation rate	0.192

Notes: This table displays the p-values for the joint placebo tests across all outcome variables of interest. The point estimates of the placebo effects are displayed in Figures B1-B5.

Table D2: Heterogeneity-Robust Estimates

	DID_ℓ estimate (1)	TWFE estimate (2)
a) Logged number of first-time URM students	0.133	0.125
	(0.0358)	(0.0419)
b) Fraction of first-time students that are URM	0.0174	0.0188
	(0.00453)	(0.00628)
c) 4-year URM grad rate	-2.32e-05	0.00247
	(0.0244)	(0.0142)
d) 6-year URM grad rate URM	-0.00507	0.00450
	(0.0255)	(0.0131)

Notes: This table reflects the DID_ℓ estimates for each outcome variable. Thus, these estimates are robust to heterogeneous and dynamic treatment effects. Column (1) contain the point estimates with bootstrapped standard errors in parenthesis. These standard errors are iterated across 100 replications. Column (2) displays the two-way fixed effects estimates from Table 3 for comparison. Clustered standard errors are in parenthesis. Sample sizes across each regression are roughly 2,820, but they somewhat vary due to some non-reporting for some left-handed variables.

E Effects of the Policy on Non-first-time students

The effects of the policy should have little impact on the enrollment outcomes of URM students that are not in their first year of college (i.e., non-freshman). This study estimates equation (1) using the outcome variables for non-freshman URM students to investigate this possibility, i.e., the logged number of non-freshman URM students and the fraction of non-freshman students that are of a URM background. The results of this exercise are provided in columns (1) and (2) of Table E1. Column (1) indicates that the policy has a negligible effect on the volume of non-freshman URM students. Interestingly, however, this policy has a small and significant impact on the fraction of non-freshmen of a URM background.

To investigate this phenomenon further, this study re-estimates equation (2) using these outcome variables to distinguish the effects between selective and highly selective colleges. The respective estimates are displayed in columns (3) and (4). It shows that the policy has little impact on the volume of non-freshman URM students at selective and highly selective colleges. Similarly, the policy has little impact on the fraction of non-freshman students from a URM background for both types of colleges. In short, there is insufficient evidence that the policy increased this fraction at each type of college. But it is possible that pooled effect on this outcome merely reflects sufficient statistical power.

Table E1: Effects of the policy on non-first-time URM enrollment

	Logged Number of non-freshmen URM students (1)	Fraction of non-freshmen that are URM (2)	Logged Number of non-freshmen URM students (3)	Fraction of non-freshman that are URM (4)
Test-optional	0.0204	0.0106*	-0.0155	0.00985
	(0.0427)	(0.00486)	(0.0705)	(0.00895)
Highly-selective $\times$ Test-optional			0.0626	2.66e-05
			(0.0866)	(0.0107)
Observations	2,824	2,819	2,824	2,819

Notes: Columns (1) and (2) displays the point estimates the policy's effects from equation (1) while columns (3) and (4) display those of equation (2). Each column corresponds to an outcome variable. The row labeled with "Test-optional" contains the point estimate for the coefficient of the treatment indicator of having the policy in place. The row labeled "Highly selective $\times$ Test-optional" contains the point estimate for the coefficient of the interaction term between the treatment indicator and another indicator for being highly selective in admissions. One star corresponds to a 5% significance level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

F Enrollment of First-time Students

This section discusses the effects of the test-optional policy on the logged number of first-time students. The purpose of this analysis is to show whether institutions within the sample are likely to implement the test-optional policy to merely boost their enrollment of students (i.e., regardless of racial background) rather than to improve their racial diversity. In this case, the policy should significantly increase the overall number of first-time students enrolling at these institutions.

Column (1) of Table F1 displays the point estimate of the test-optional policy using equation (1). It suggests that within the sample as a whole, the policy has a small, positive impact on the volume of first-time students enrolling at these colleges (i.e., by 1.33%). However, this point estimate is insignificant. Column (2) displays the point estimates from estimating equation (2) to distinguish the enrollment effects across institution selectivity, and it shows that selective colleges increased their first-year enrollment by 2.83% as a result of the policy. This point estimate is also insignificant. Similarly, the policy has little impact on this outcome for highly selective colleges ($p = .889$). Finally, column (3) displays the point estimates from estimating equation (3) to distinguish colleges adopting the policy through versus after 2007. These estimates show that the two groups experienced negligible and indistinguishable effects on overall first-year enrollment ($p = .559$). All in all, these results suggest that most of the colleges in the sample are likely to be aiming to improve campus racial diversity rather than simply bolstering their enrollment levels.

Table F1: Effect of Test-Optional Policy on Logged First-time Enrollment

	Main Specification (1)	Selectivity Heterogeneity (2)	Adoption Timing (3)
Test-optional	0.0133	0.0283	
	(0.00955)	(0.0167)	
Highly Selective × Test-optional		-0.0267	
		(0.0204)	
Early adopter			0.0195
			(0.0159)
Late adopter			0.00914
			(0.0104)
p-value			0.539
Observations	2,820	2,820	2,820

Notes: This table displays the results from estimating equation (1), (2), and (3) on the logged number of first-time students. The point estimates from each equation correspond to columns (1)-(3), respectively. The row labeled with "Test-optional" contains the point estimate for the coefficient of the treatment indicator of having the policy in place. The row labeled "Highly selective × Test-optional" contains the point estimate for the coefficient of the interaction term between the treatment indicator and another indicator for being highly selective in admissions. The row labeled "Early adopter" corresponds to the estimated effect of the policy on institutions that dropped the test requirement early. Similarly, the row labeled "Late adopter" corresponds to the effect on institutions adopting the policy later. Standard errors are in parenthesis and clustered by institution.

G Inclusion of HBCU Institutions

The initial sample of institutions included three Historically Black Colleges & Universities (HB-CUs): Morehouse College, Spelman College, and Fisk University. All of these are considered to be "selective" by the USNWR. However, they are excluded from the primary analyses. Tables G1 and G2 replicate the results from Tables 3 and 4 while including HBCUs. In Table G1, the point estimates for the effects of policy on URM enrollment outcomes (e.g., the fraction of freshman students that are URM) are slightly larger than the point estimates from Table 3. Similarly, in Table G2, the point estimates for these outcomes among selective colleges are also larger than the point estimates from Table 4. The larger point estimates could be attributed to the trends in these outcomes that these HBCU institutions follow across time within the panel. However, the difference in point estimates between selective and highly selective colleges remains statistically insignificant.

Table G1: Primary Results (with HBCUs)

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	6-year URM graduation rate (4)
Test-optional	0.140**	0.0218**	0.00123	0.00609
	(0.0428)	(0.00667)	(0.0140)	(0.0129)
Observations	2,877	2,881	2,837	2,869

Notes: This table displays the results from estimating equation (1) on all outcome variables. This table, however, reflect the point estimates from a sample that includes HBCUs. The row labeled with "Test-optional" contains the point estimate of the TWFE coefficient. Each column correspond to an outcome variable. Columns 4-7 correspond to graduation outcomes, so the regressions for those use a lagged treatment indicator (i.e., $P_{i,t-6}$). One stars indicates a 5% significance level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression may vary slightly due to some non-reporting for some left-handed variables. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

Table G2: Effects of the Policy across Selectivity (with HBCUs)

	Logged Number of first-time URM students (1)	Fraction of first-time students that are URM (2)	4-year URM graduation rate (3)	6-year URM graduation rate (4)
Test-optional	0.199**	0.0335*	-0.0508*	-0.0461*
	(0.0727)	(0.0133)	(0.0204)	(0.0196)
Highly selective × Test-optional	-0.0785	-0.0155	0.0656*	0.0632**
	(0.0872)	(0.0151)	(0.0252)	(0.0236)
Observations	2,877	2,881 2,837	2,869	

Notes: This table displays the point estimate from equation (2), which distinguishes institutions by their degree of admission selectivity. However, these reflect a sample that includes HBCUs. The row labeled with "Test-optional" contains the point estimate for the coefficient of the treatment indicator of having the policy in place. The row labeled "Highly selective × Test-optional" contains the point estimate for the coefficient of the interaction term between the treatment indicator and another indicator for being highly selective in admissions. Each column corresponds to an outcome variable. The point estimates for the coefficients of the control variables are not reported in this table. One star indicates a 5% significance level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression slightly vary due to non-reporting for some left-handed variables.

H Non-URM Graduation Rates

In this section, this study investigates whether the policy affects the graduation rates for non-URM students (e.g., White and Asian students). Columns (1) and (2) of Table H1 shows that the policy has little impact on 4 and 6-year graduation rates for non-URM students throughout the sample as a whole. Furthermore, columns (3) and (4) indicate that neither selective nor highly selective colleges experience non-URM graduation rate effects as a result of this policy. This shows that the decline in the URM graduation rates at selective institutions are not driven by a drop in the overall campus graduation rates.

Table H1: Graduation rate of non-URM students

	4-year grad rate (1)	6-year grad rate (2)	4-year grad rate (3)	6-year grad rate (4)
Test-optional	0.000254	0.00689	-0.0187	-0.00582
	(0.00793)	(0.00701)	(0.0192)	(0.0148)
Highly selective × Test-optional			0.0253	0.0148
			(0.0210)	(0.0169)
Observations	2,792	2,821	2,792	2,821

Notes: Columns (1) and (2) of this table results from estimating equation (1). Similarly, columns (3) and (4) results from estimating equation (2). Each column corresponds to an outcome variable. One stars indicates a 5% significance level and two stars indicate a 1% level. Standard errors are in parenthesis and clustered by institution. Sample sizes across each regression may vary slightly due to some non-reporting for some left-handed variables.

I Analysis at the State Level

Colleges included within the sample are located across 37 states. Of these, 29 contain at least one test-optional institution. However, some states, such as Pennsylvania, have multiple test-optional colleges. As discussed in Section 5.1, colleges may drop their test requirement when neighboring colleges within a state do so. For example, as shown in Table A1, Southwestern University in Texas dropped the test requirement one year after Austin College. Furthermore, several colleges from the same state switched to test-optional admissions simultaneously (e.g., Trinity College and Wesleyan University in 2015).

This study collapses the panel to the state level to investigate this possible phenomenon. It estimates a version of equation (1) displayed below:

$$y_{st} = \beta A_{st} + \boldsymbol{X'}_{st}\gamma + \theta_s + \lambda_t + \eta_{st}. \tag{5}$$

Here, y_{st} either corresponds to the logged number of URM students in each state s or the share of students in each state from a URM background (i.e., these are state-level freshman URM enrollment outcomes). A_{st} corresponds to the share of students in each state that are enrolled at a test-optional college. Finally, θ_s and η_{st} are the state fixed effects and the stochastic error term, respectively. The covariates within $\boldsymbol{X}_{st}$ are primarily state-level averages (e.g., logged average tuition & fees among colleges in state s).

The results of estimating equation (5) are displayed in Table I1. The point estimates for each state-level freshman URM enrollment outcome are, in short, noisy. Thus, there is little evidence that colleges self-select themselves for treatment to follow the steps of other colleges in the same state.

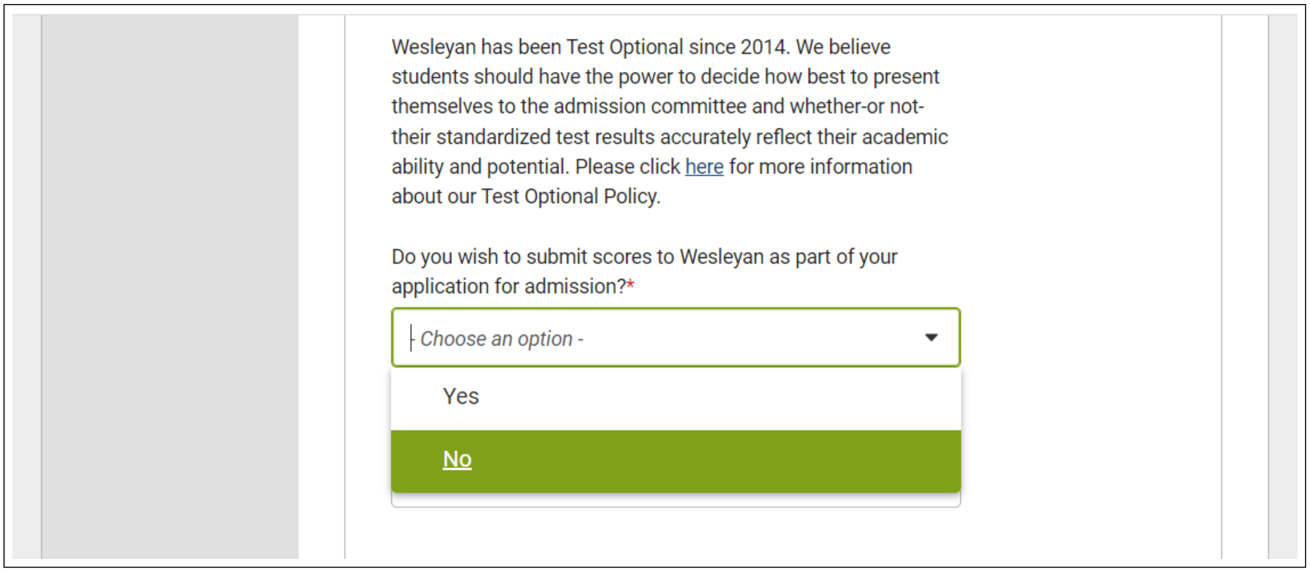
Table I1: State-level Effects of Test-Optional Admissions

	Logged number of URM (1)	Portion of first-time that are URM (2)
Share Test-Optional	0.115	0.00921
	(0.110)	(0.0148)
Observations	703	703

Notes: This table results from estimating equation (5). Each column corresponds to an outcome variable related to freshman URM enrollment. Standard errors are in parenthesis and clustered by state.

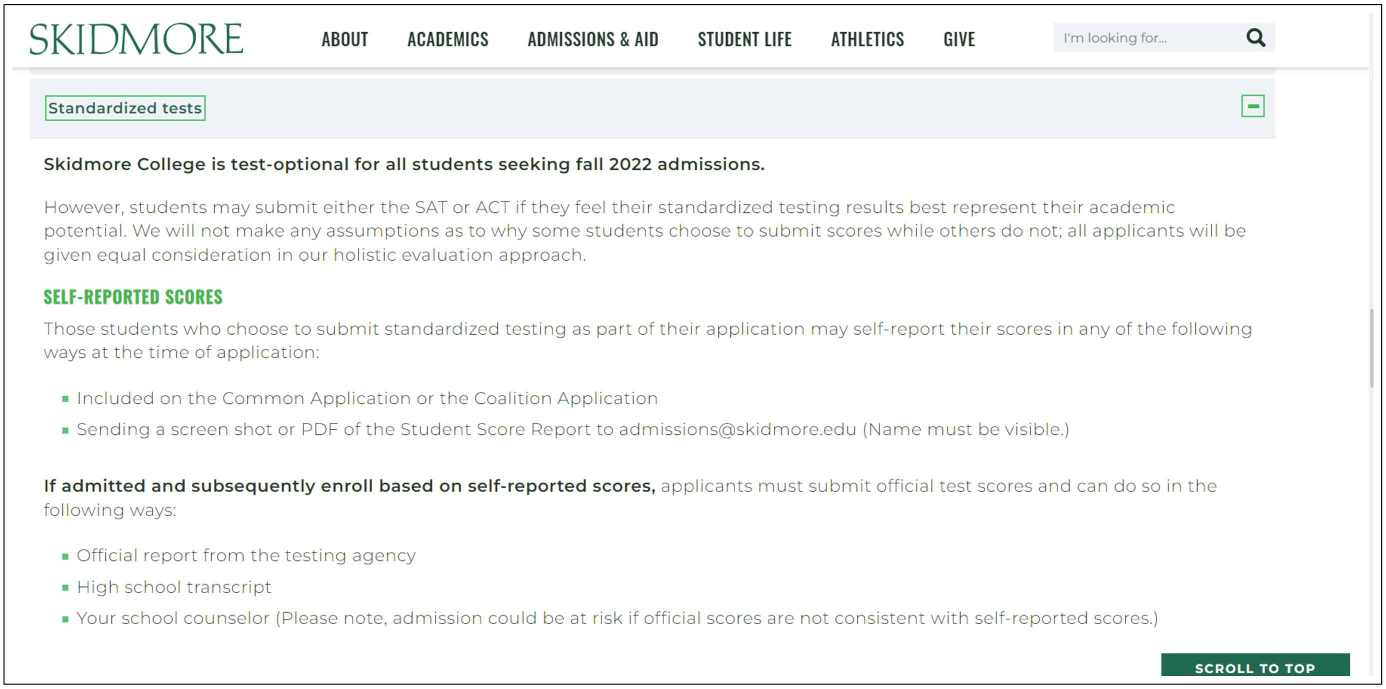
J Miscellaneous Figures and Tables

Figure J1



Notes: This figure is a screenshot of the Common Application page for Wesleyan University, a test-optional college. Similar to other test-optional institutions, the application for this college asks the applicant to choose whether they wish to submit their SAT or ACT test scores for admission consideration.

Figure J3


SKIDMORE

ABOUT ACADEMICS ADMISSIONS & AID STUDENT LIFE ATHLETICS GIVE

I'm looking for...

Standardized tests

Skidmore College is test-optional for all students seeking fall 2022 admissions.

However, students may submit either the SAT or ACT if they feel their standardized testing results best represent their academic potential. We will not make any assumptions as to why some students choose to submit scores while others do not; all applicants will be given equal consideration in our holistic evaluation approach.

SELF-REPORTED SCORES

Those students who choose to submit standardized testing as part of their application may self-report their scores in any of the following ways at the time of application:

- Included on the Common Application or the Coalition Application
- Sending a screen shot or PDF of the Student Score Report to admissions@skidmore.edu (Name must be visible.)

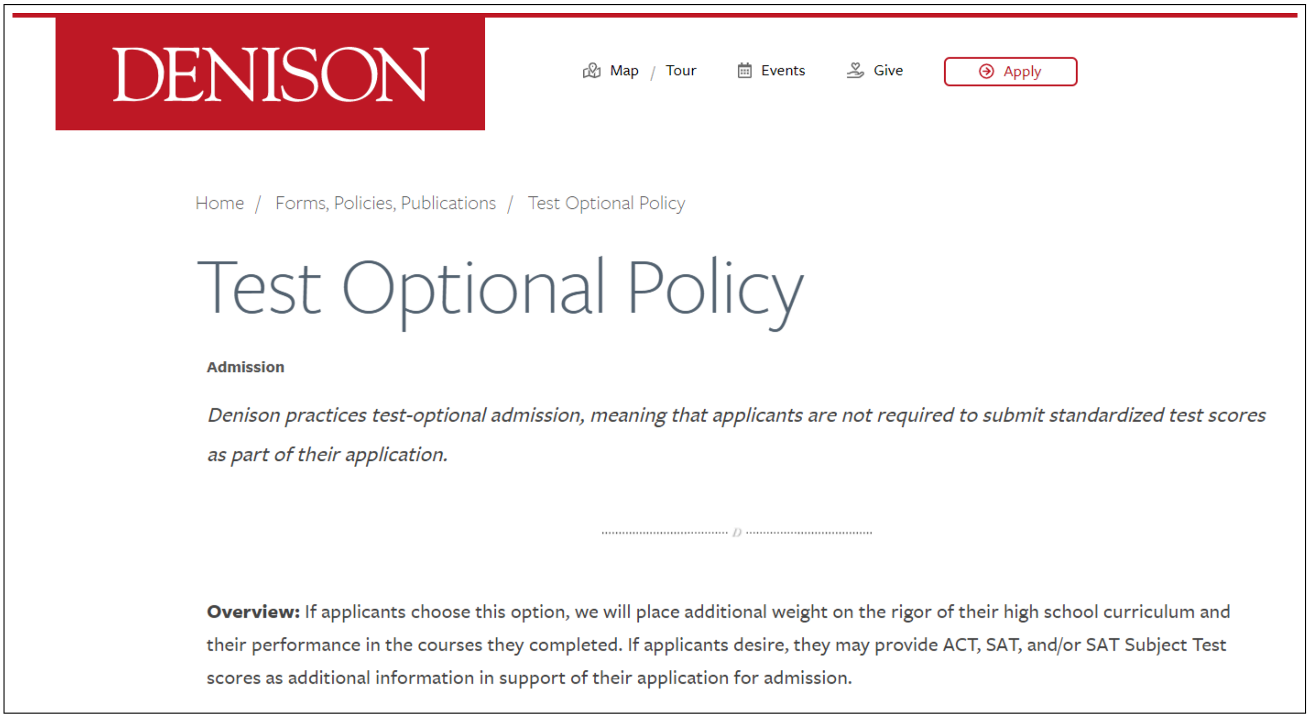
If admitted and subsequently enroll based on self-reported scores, applicants must submit official test scores and can do so in the following ways:

- Official report from the testing agency
- High school transcript
- Your school counselor (Please note, admission could be at risk if official scores are not consistent with self-reported scores.)

SCROLL TO TOP

Notes: This figure is a screenshot of the admissions page for a test-optional school, Skidmore College. They suggest that they will provide equitable consideration between applicants that choose to submit their SAT or ACT test scores and those that do not.

Figure J3



Notes: This figure is a screenshot of the admissions page for a test-optional school, Denison University. They indicate that applicants withholding their standardized test scores will be scrutinized more heavily by the rigor of their high school coursework, as well as their performance in their respective classes.