

Untested Admissions: Examining Changes in Application Behaviors and Student Demographics Under Test-Optional Policies

Christopher T. Bennett 
Vanderbilt University

This study examines a diverse set of nearly 100 private institutions that adopted test-optional undergraduate admissions policies between 2005–2006 and 2015–2016. Using comparative interrupted time series analysis and difference-in-differences with matching, I find that test-optional policies were associated with a 3% to 4% increase in Pell Grant recipients, a 10% to 12% increase in first-time students from underrepresented racial/ethnic backgrounds, and a 6% to 8% increase in first-time enrollment of women. Overall, I do not detect clear evidence of changes in application volume or yield rate. Subgroup analyses suggest that these patterns were generally similar for both the more selective and the less selective institutions examined. These findings provide evidence regarding the potential—and the limitations—of using test-optional policies to improve equity in admissions.

KEYWORDS: higher education, test-optional, college admissions, standardized tests

Introduction

Throughout the 20th century, selective 4-year colleges and universities in the United States widely adopted standardized tests as admissions requirements (Syverson, 2007). Despite early aspirations that such tests could serve as a “census of human abilities” capable of reliably identifying talented individuals from a variety of backgrounds (Carmichael & Mead, 1951, p. 196), there remain strong correlations between college admissions test scores and race, gender, and socioeconomic status (e.g., Bowen et al., 2009). Alon and Tienda (2007)

CHRISTOPHER T. BENNETT is a PhD candidate in the Department of Leadership, Policy, and Organizations at Vanderbilt University, 230 Appleton Place, PMB #414, Nashville, TN 37203, USA; e-mail: chris.bennett@vanderbilt.edu. His research focuses on topics related to college access, graduate/professional education, and labor market outcomes for students.

have argued that the “apparent tension between merit and diversity exists only when merit is narrowly defined by test scores” (p. 487). Amid such concerns, over 10% of selective institutions shifted to test-optional admissions policies between 1987 and 2015 (Furuta, 2017). Still more institutions have adopted test-optional policies, at least temporarily, following large-scale testing cancellations in the wake of the coronavirus outbreak in 2020 (Anderson, 2020).

Given the increasingly widespread adoption of test-optional admissions strategies and the significant benefits that accrue to individuals who attend selective institutions (e.g., Bowen & Bok, 1998), the impact of such test-optional policies warrants further attention. Previous work on test-optional policies has largely focused on selective liberal arts colleges, which accounted for nearly all of the earliest adopters (Belasco et al., 2015). Considerably less attention has been devoted to the implications of the test-optional movement as it continued to expand beyond selective liberal arts colleges. The present study offers insights into this much wider pool of test-optional institutions, encompassing a range of selectivity levels and institution types. In doing so, this study takes extensive steps to identify suitable comparison groups, relies on policy enactment dates corroborated from multiple sources, and follows outcome measures for more years postenactment than other studies, on average.

To provide greater understanding of the implications of test-optional policies, the primary analysis for this study compares nearly 100 private institutions that implemented test-optional policies between 2005–2006 and 2015–2016 (a treated group of earlier-adopters) to more than 100 others that subsequently enacted or announced test-optional policies by December 2019 (a comparison group of later-adopters). In addition, due to the variety of institutions that adopted test-optional policies within this timeframe, I use three matching methods to identify institutions that closely resemble one another on key observable dimensions. This approach enables me to emphasize results that are consistent across multiple comparison groups, suggesting they are not contingent on the exact matching approach used.

Using comparative interrupted time series (CITS) and difference-in-differences (DD) analytic strategies, I assess the effects of test-optional policy enactment on measures related to undergraduate admissions and the composition of the student body. Relying on the variation in policy adoption timing for more test-optional institutions than any prior published study, I find that the enrollment of Pell Grant recipients, underrepresented racially/ethnically minoritized (URM) students, and women increased following test-optional policy enactment, relative to trends at matched peer institutions. I do not find clear overall evidence of an increase in applications, though there may have been a short-term rise during the first few years of the policy. The study detects no consistent relationships between test-optional policies and an institution’s yield rate, although it is not possible to entirely rule out modest effects. These findings are robust to a variety of alternative specifications that produce qualitatively similar estimates. Subgroup analyses suggest that

more selective and less selective institutions alike experienced the increases in Pell Grant recipients, URM students, and women.

Background

To contextualize this study, it is worthwhile to consider the role of standardized tests in college admissions and previous research on test-optional policies. I conclude this section with an overview of the conceptual framework for the study, including a discussion of the roles that students and institutions play in shaping admissions and enrollment outcomes.

Test-Optional Policy Adoption

Prior to the advent of standardized testing in the early 20th century, American universities typically admitted students through institution-specific examinations or guaranteed admission for graduates of preapproved high schools (Wechsler, 1977). Descended from IQ tests and Army Alpha tests used to assess military recruits, the SAT emerged in 1926 as an alternative mechanism for assessing applicants (Zwick, 2004). Among the earliest uses of the SAT was as a selection criterion for a national scholarship program at Harvard University, an initiative through which Harvard president James Bryant Conant sought to expand access to students from a wider array of socioeconomic backgrounds (Karabel, 2006). In part through Conant's advocacy for the SAT, other selective institutions began adopting the SAT in their admissions practices, laying the groundwork for a long-standing link between standardized testing and selective college admissions (Zwick, 2019). With an expanding pool of college applicants as a result of the G.I. Bill, institutions increasingly turned to the SAT to distinguish between candidates (Lemann, 1999). By 1959, a large-scale competitor, the ACT, had emerged, as well.

As the pool of college students has expanded in subsequent decades, so too have debates about the role these standardized test scores play in the admissions process. One primary set of concerns centers on persistent disparities in standardized test scores across various groups. In particular, average SAT and ACT scores vary by a student's socioeconomic status and racial/ethnic group (Dixon-Román et al., 2013). For example, among high school seniors in 2010, Asian students averaged a combined 1,110 points on the Critical Reading and Mathematics sections, compared to 857 for Black students, 906 to 921 for subgroups of Hispanic students, 977 for Native American students, and 1,064 for White students (College Board, 2010). Furthermore, on average, women underperform on the SAT relative to men, especially when considering performance in high school courses (Leonard & Jiang, 1999). For college-bound seniors who took the SAT in 2010, for instance, men averaged 34 points higher on the Mathematics section and 5 points higher on the Critical Reading section (College Board, 2010).

These differences do not originate with college admissions tests but are instead reflective of long-standing, systemic variation in educational resources and opportunities that also manifest themselves in achievement disparities on standardized tests at the K–12 level (Duncan & Magnuson, 2005; Reardon et al., 2019). In addition to such underlying forces, differential access to test preparation services may partially account for such score gaps, particularly for students from low-income backgrounds (Buchmann et al., 2010). Other proposed explanatory factors include stereotype threat, in which negative stereotypes lead members of the stereotyped group to underperform (Walton & Spencer, 2009), and content that advantages members of certain groups (Rosner, 2012). Regardless of the origin, the strong correlations between background characteristics and test scores could lead applicants to believe that standardized tests reinforce rather than displace barriers to social mobility.

A second group of concerns regarding standardized testing involves the predictive validity of scores. Much of the predictive validity literature has found that standardized test scores offer incremental improvements in predicting college grades and graduation rates, relative to high school grades alone (e.g., Westrick, et al., 2015). However, other work has conveyed the importance of considering heterogeneity in the predictive validity for the SAT and ACT. For instance, using data from 2006 test-takers, Mattern et al. (2008) found variation in the relationship between SAT scores and first-year college grade point average (FYGPA) by gender, race/ethnicity, and a student's best language. They determined that SAT scores overpredicted FYGPA for men by 0.15 standard deviations while underpredicting women's FYGPA by 0.13 standard deviations. Across racial/ethnic groups, the relationship between SAT scores and FYGPA ranged from an overprediction of 0.22 standard deviations (for Native American students) to an underprediction of 0.03 standard deviations (for White students). Meanwhile, for students who reported speaking another language better than English, the SAT underpredicted FYGPA by 0.33 standard deviations. More recent evidence has also highlighted the importance of considering high school context when evaluating the relationship between ACT scores and college graduation rates, with Allensworth and Clark (2020) finding that "the signal provided by ACT scores is much smaller than the noise introduced by school effects" (p. 209). In light of such variation, test scores' utility varies depending on the purpose for which they are used, and the decision of how much weight to assign to test scores is consequential (Bowen et al., 2018).

At least partly in response to such concerns, a growing number of institutions have chosen to adopt test-optional admission policies (Furuta, 2017). These policies permit applicants to decide whether to submit standardized test scores in their application. Admissions professionals at test-optional institutions then use the available evidence to determine which applicants to admit. Test-optional institutions state that they do not penalize applicants who choose to omit test scores (Syverson et al., 2018). For students who do not submit test scores, institutions must rely more heavily on other factors when making an

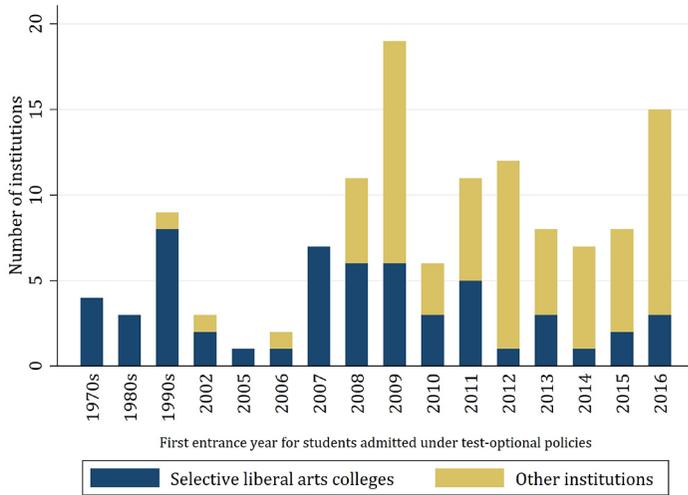


Figure 1. Number of institutions enacting test-optional admissions over time, by institution type.

Note. Prior to 2001–2002, policy enactment is listed by decade.

admission decision (e.g., grades, extracurriculars), although the precise balance is unknown. In some cases, test-optional institutions require additional application materials (e.g., supplementary essays) in place of test scores, or they may require test scores for placement purposes once a student enrolls.

While Bowdoin College and a handful of other institutions have maintained test-optional policies for about half a century, the preponderance of test-optional policy adoption has taken place within the past two decades (Furuta, 2017). Early on, test-optional policies were overwhelmingly concentrated in selective liberal arts colleges. By the 2010s, however, the test-optional movement expanded to reach a variety of other sectors. In fact, Figure 1 illustrates that selective liberal arts colleges account for the minority of institutions adopting test-optional policies in more recent years. For instance, of private institutions that first enacted test-optional policies for students entering between 2008–2009 and 2015–2016, only one quarter were liberal arts colleges. Despite such diversification in terms of the Carnegie classifications, test-optional adoption in this period was overwhelmingly limited to private institutions, which are therefore the focus of this study.

Impacts of Test-Optional Policy Adoption

Before COVID-19 dramatically constrained the availability of standardized testing, institutions espoused two primary rationales for adopting test-

optional policies. First, institutions frequently invoked the notion that a person is “more than just a number,” arguing a test-optional policy is consistent with a holistic admissions approach and the university’s values more broadly (e.g., Hanover College, 2019). Second, they commonly explained that they expected test-optional policies to increase student body diversity along several dimensions. For instance, the dean of admissions at Bowdoin College suggested their use of the policy increased “geographic, socioeconomic, racial, [and] ethnic” diversity (Goldfine, 2017). Other case studies have similarly reported that student body diversity increased after test-optional policy adoption, though they do not claim their findings represent average effects for all test-optional adopters (e.g., Syverson et al., 2018). Early data simulations also suggested that improved socioeconomic and racial/ethnic diversity would be likely outcomes of the policies (Espenshade & Chung, 2010).

A limited body of research has sought to identify the causal effects of test-optional policies. Focusing on 32 test-optional selective liberal arts colleges, Belasco et al. (2015) used a DD analytic strategy and determined that the policies were more effective at achieving latent, rather than the manifest, aims. Specifically, they found that the policies had no significant impact on the proportion of URM students or Pell Grant recipients enrolled, although test-optional selective liberal arts colleges experienced increased application volumes and higher reported SAT scores after implementing the policy. Such findings suggest that the implementation of test-optional policies at a small group of early adopters may have enhanced the perceived selectivity of the institution without meaningfully improving the socioeconomic and racial diversity on those campuses. Analyzing the same set of institutions with a DD approach combined with propensity score matching (PSM), Sweitzer et al. (2018) obtained generally similar results, finding that test-optional policies resulted in increased SAT scores but had no statistically significant effects on applicant volume, acceptance rates, or the number of URM students. Subsequent work from Rosinger and Ford (2019) analyzing a similar group of test-optional liberal arts colleges found no significant evidence of changes in enrollment by family income quintile.

Looking beyond selective liberal arts colleges, Saboe and Terrizzi (2019) used a DD approach to examine effects of test-optional policies at 4-year institutions of all types, finding few effects apart from a short-term initial increase in applications. There are several caveats to the analysis in that article, however. For instance, Saboe and Terrizzi compared test-optional institutions to all 4-year test-requiring institutions, regardless of other institutional characteristics such as selectivity. Since many 4-year institutions are actually or effectively open-access, their admissions strategies and goals are likely to differ from the selective institutions that adopt test-optional policies. Additionally, Saboe and Terrizzi’s measure of racial diversity was a count of all non-White students, rather than focusing on historically underrepresented students in particular. Finally, the data analyzed were confined to the period

between 2009 and 2014, a range of years immediately following the Great Recession that does not include dozens of test-optional adopters outside that time period.

Overall, then, prior literature finds few effects of test-optional policies, apart from occasional evidence of increases in applications and reported standardized test scores. However, as outlined above, that prior literature either focused exclusively on selective liberal arts colleges or compared test-optional adopters to a counterfactual set that likely differed from the test-optional adopters in meaningful ways. In addition, the analytic time period in prior studies did not allow the majority of institutions to be followed for 4 years after test-optional policy adoption. Such a comparatively limited observation window makes it difficult to interpret estimates from prior research regarding outcomes such as Pell Grant recipient enrollment, which takes 4 years to be fully evident.

Building on previous literature, this study identifies suggestive evidence for the effects of test-optional policy adoption at a broad array of private institutions. To do so, I examine a larger group of test-optional institutions than prior studies, allowing me to evaluate the policy at the diverse coalition of institutions that the test-optional movement has come to encompass. I also follow outcomes for more years postadoption, on average, facilitating analysis of long-term outcomes. As a result, this study provides some of the most recent, largest scale evidence on the effects of test-optional policies for enrollment management metrics and student body diversity at private institutions in the United States.

Conceptual Framework

The underlying process that connects test-optional policy implementation and student behavior is primarily student decision-making regarding whether and where to enroll in higher education. To better understand how students make these decisions, I draw on Perna's (2006) model of college choice. At the same time that students undertake the college choice process, of course, institutions themselves are engaged in an effort to identify and attract applicants who are well-suited for the institution's aims. Therefore, while this conceptual framework emphasizes the student's role in choosing where to apply and ultimately matriculate, it also addresses some institutional actions that can facilitate or constrain student decision-making.

Seeking to explain differential college choice processes across a variety of groups, Perna (2006) introduced a conceptual model of college choice that integrated considerations from economic and sociological traditions. Perna's model identified college choice in part as a human capital investment decision, with students evaluating the expected costs and benefits of their college choice in light of their own academic background and family resources. In Perna's conceptualization, this human capital decision is nested within four layers of context. The first

layer involves the interrelationship between a student's demographic characteristics (e.g., race/ethnicity, gender), cultural capital, and social capital. Together, these attributes help form a student's *habitus*, an internalized set of dispositions and preferences that undergirds an individual's behaviors and decision-making (Bourdieu, 1977). The second layer covers a student's high school and community context, including the resources available to facilitate or inhibit college choice activities. A third layer of context addresses the role of institutions of higher education, including their recruitment efforts, geographic proximity, and other attributes. The fourth and final layer extends to the larger social, economic, and political context in which a student is engaging in the college choice process. Thus, under Perna's model, the college choice process is ultimately a human capital investment decision, though students' assessments differ based on several layers of context within which they are embedded.

In what ways do these models, in combination with prior research, suggest test-optional admissions practices could alter the college choice process? To answer this question, I outline a theory of change that uses Perna's (2006) college choice model to identify several potential channels by which adoption of test-optional policies could affect students' application decisions. I focus on four potential channels: increased publicity, removal of application barriers, shifts in students' perceptions of academic fit, and alignment between students' ideals and institutions' professed values. From there, I discuss the ways that test-optional policy adoption could affect admissions professionals' and students' decisions (to admit and to enroll, respectively), jointly resulting in student body composition shifts.

First, the policy change may increase a student's awareness of a test-optional institution, a necessary precursor to becoming part of a student's choice set. Since most selective colleges and universities required test scores during the analytic time period, newly test-optional institutions regularly received media attention. Such media exposure can prove valuable in the competitive market for private colleges and universities, many of which have comparatively small student bodies and are unable to rely on the same name recognition as public flagships, for example. This increased publicity could be relevant for students with various levels of prior access to social capital. For instance, information on admissions policy changes may accrue to students who can rely on the expertise of college-educated parents in the college search process (Engle & Tinto, 2008), students embedded in high school contexts with greater college-going cultures (K. J. Robinson & Roksa, 2016), and wealthier students with greater access to college and private counselors (McDonough et al., 1997; Plank & Jordan, 2001). Conversely, increased publicity may be most salient for students with less prior access to information about the universe of college options.

Second, by permitting students to apply without standardized test scores, test-optional policies may reduce impediments to applying (e.g., structural barriers in the community context, such as limited transportation to testing

sites), thereby making such institutions more attractive. In light of differential SAT- and ACT-taking rates across student groups (e.g., Klasik, 2012), individuals for whom test-taking represents an application barrier appear disproportionately likely to be URM students and students from low-income backgrounds. These differences in test-taking rates may originate from a number of factors. For instance, prior research has shown shifts in college entrance exam completion based on the availability of test-taking sites, particularly for students at schools with high proportions of minoritized and low-income students (Bulman, 2015). Seeking to reduce such gaps in test-taking, a growing number of states have begun initiatives to promote universal participation in college entrance examinations (e.g., Goodman, 2016; Hyman, 2017). Even for students who take the SAT or ACT, however, research from Pallais (2015) suggested that other seemingly small financial and behavioral hurdles, such as the default number of free score reports available, can affect the types of colleges where low-income students apply and enroll. From this perspective, then, the availability of test-optional admissions policies could prove attractive to URM students and students from low-income backgrounds.

Third, test-optional policies could shift students' perceptions of their academic fit for an institution. Given that students consider likely admissions outcomes when identifying institutions of interest (Manski & Wise, 1983), some individuals may be deterred from applying to selective institutions because of a (potentially inaccurate) expectation that their standardized test scores would not make them competitive candidates. Based on average standardized test scores, such concerns might be especially pronounced among (but by no means unique to) students from low-income backgrounds, URM students, and women—particularly if they are embedded in school and community contexts in which few students apply to selective private institutions. By adopting a test-optional admissions policy, selective private institutions may prompt students to reevaluate their likelihood of acceptance, especially if a student's class rank and GPA are more typical of admitted students than their standardized test scores. Such reevaluations may be common among women, who are more likely than men to take themselves out of the running for jobs due to concerns that they do not meet necessary qualifications (Mohr, 2014). If test-optional policies shift women's perceptions of necessary qualifications for admission to ones that prioritize GPAs, on which they outperform men (Niederle & Vesterlund, 2010), they may result in increased applications from women.

Also supporting the potential role of this third channel, literature on academic “undermatching” suggests low-income students do not apply to or matriculate at selective institutions at the same rates as higher income peers, even when they possess academic credentials typical of admitted students (e.g., Dillon & Smith, 2017; Hoxby & Avery, 2012). A substantial number of high-ability students also underestimate their academic competitiveness to such an extent that they fail to take standardized tests altogether

(Goodman, 2016), foreclosing any possibility of enrolling at a selective test-requiring institution. Evidence from Texas's Top Ten Percent Plan demonstrates that substantial shares of students (especially those who are Black or Hispanic) "undermatch" despite having guaranteed admission to selective institutions based on class rank (Black et al., 2015). Consequently, it is clear there are constraints on the extent to which applicants shift their application behaviors even when perfectly informed about the likelihood of acceptance.

A fourth channel by which test-optional policies might shift students' application behaviors is by changing the perceived alignment between the values of the student and the institution. For instance, some students may interpret test-optional policies as signals that an institution values individuality and considers a student's unique needs and contributions. Empirical support for such a relationship comes from Furuta (2017), who found increased odds of test-optional policy adoption among institutions with expanded notions of "student personhood" (e.g., availability of self-designed majors), even after controlling for a variety of factors. This channel may operate through either the higher education layer of context or the broader social, economic, and policy context. Within the higher education context, institutions themselves may strategically deploy their test-optional status to assist in recruitment efforts. Alternatively, students' evaluations of institutions with test-optional policies could shift based on broader societal narratives regarding the appropriate levels of standardized testing (e.g., Zernike, 2015) and evolving conceptions of what constitutes equitable notions of merit (Warikoo, 2016).

Ultimately, the extent to which shifts in application behaviors lead to differences in enrollment depends on the decisions of both admissions officers and admitted students. Test-optional institutions typically claim to evaluate students on equal footing regardless of whether they submit test scores, although there is little existing literature on the exact processes by which admissions officers evaluate candidates that differ by test score availability (for a case study, see M. Robinson & Monks, 2005). Even assuming that admissions officers do not penalize score nonsubmitters, one possibility is that test-optional policies primarily increase applications from students whom admissions officers are unlikely to admit. Such a scenario could result in improvements in selectivity metrics (e.g., application volume, reported standardized test scores) without meaningfully changing the student body composition, similar to earlier findings from Belasco et al. (2015). Alternatively, the test-optional policy could shift the higher education context level of Perna's (2006) college choice model and provide institutions with greater latitude to admit students who present compelling cases for admission but whose standardized test scores (or lack thereof) would have previously served as a barrier to admission. In the latter case, accepted students may exhibit greater

diversity along several dimensions, while the institution is able to maintain or increase its reported standardized test scores.

The test-optional policy introduction could also have implications for the matriculation decisions of admitted students, thereby affecting the yield rate (i.e., share of admitted students who enroll). Consider the case of a student with a compelling application, apart from a low SAT score. At a highly selective institution with test-optional admissions, they might have received strong consideration as a test score nonsubmitter. In contrast, highly selective institutions that required test scores may have been less likely to admit the student. In such cases, the test-optional institution may have represented the most selective institution to which a student received admission, which would have factored into their human capital decision. If the test-optional policy facilitated a sufficiently large number of matches to students for whom they represented the preferred choice, an increase in the yield rate would be evident. A higher yield rate may also be apparent if students believed test-optional institutions were more aligned with their values, as suggested in the fourth proposed channel discussed above. In an era when the number of college applications per student increased substantially (Clinedinst & Koranteng, 2017), increases in yield rate can help reduce uncertainty for institutions in the admissions process, with meaningful implications for university budgeting and planning.

To explore the extent to which test-optional policies affect subsequent application and enrollment behaviors, I therefore address two primary research questions.

Research Question 1: Following enactment of test-optional admissions, do institutions experience significant changes in key measures of application behavior (i.e., number of applications and yield rate) and student demographics (i.e., Pell Grant recipients, URM students, women)?

Research Question 2: How do the effects of test-optional policies on application behaviors and student demographics vary based on an institution selectivity and institution type?

Data and Measures

To address the research questions, I assembled data from multiple sources, resulting in an institution-level data set of private 4-year institutions covering the academic years 2001–2002 through 2015–2016. Using a multistep process to verify the adoption period at each institution, I identified 99 private institutions that enacted test-optional admissions policies for students entering between 2005–2006 and 2015–2016 (“earlier-adopters”). As a comparison group, I also identified an additional 118 institutions that had enacted or announced test-optional policies for academic years 2016–2017 or later, as

of December 2019 (“later-adopters”). I excluded institutions that did not have selective admissions according to the 2003 Barron’s competitiveness index (i.e., not at least “less competitive”), as well as institutions that did not award bachelor’s degrees or were ineligible for federal Title IV aid. Additionally, I excluded institutions designated as “specialized institutions” in the Carnegie classifications, which include art schools and other programs for which standardized test scores may play a more limited role in admissions. The methodology for identifying test-optional institutions is detailed below in the section titled “Classifying Treatment Institutions.” See Supplemental Appendix Tables 1 and 2 for the full lists of earlier- and later adopting institutions (available in the online version of the journal).

Outcome Measures

Aligned with the conceptual framework, the outcome measures include overall applications, overall yield rate, and enrollment among certain groups of students. The primary source for institution-level variables was the Integrated Postsecondary Education Data System (IPEDS). Outcome variables from IPEDS included the number of first-year applications, yield rate (i.e., entering students divided by accepted students), first-time full-time (FTFT) URM students, and FTFT women.¹ (For consistency, I linked application data based on entry year rather than application year.) While some prior research on test-optional policies also explored standardized test scores as an outcome measure, approximately two fifths of test-optional institutions no longer reported standardized test scores to IPEDS by 2016. Thus, out of concern about the extent of data missingness for standardized test scores and the likely bias in reported scores, this study does not focus on standardized test scores as an outcome.

Data on the total number of Pell Grant recipients were available through the Office of Postsecondary Education (U.S. Department of Education, 2018). Although Pell Grant status is an imperfect proxy for students from lower income backgrounds (Rosinger & Ford, 2019; Tebbs & Turner, 2005), its ubiquity provides incentives for institutions to increase enrollment of Pell-eligible students in particular if they are indeed expanding access (Hoxby & Turner, 2019).

Control Measures

Additional institutional attributes from IPEDS that measure key sources of differentiation between the treatment and comparison groups include institutional Carnegie classification and geographic region, which prior research identified as associated with test-optional policy adoption (Furuta, 2017). Through the Delta Cost Project, I acquired four more control variables, which were standardized IPEDS variables (Hurlburt et al., 2017). One such measure was the full-time equivalent (FTE) number of students, which represents a rough indication of the overall size and complexity of the organization. Instructional expenditures per FTE serve as a proxy for resources that affect

undergraduate academic experiences. I also included student services expenditures per FTE to help account for differences in institutional outlays for items such as student activities, admissions, and recruitment, which may help shape prospective applicants' perceptions of campus life. Last, based on student responsiveness to the sticker price of college (e.g., Hoxby & Turner, 2015), I also controlled for listed tuition and fees.

Since the second research question focused on variation in effects by an institution's selectivity level, I also incorporated the Barron's competitiveness index as a measure of institutional selectivity. I used the 2003 Barron's competitiveness index for all institutions, which considers acceptance rate, class rank of admitted students, and test scores of admitted students in its calculation (Barron's Educational Series, 2002). Conceptually, prior research has used a variety of approaches to grouping the categories of the Barron's competitiveness index, so there is no clear standard for grouping (e.g., Braxton, 1993; Light & Strayer, 2000). In the current study, I categorized those institutions designated as "most competitive" or "very competitive" as "more selective," and remaining institutions as "less selective."

Finally, I also included an indicator of whether an institution had a "no-loan" or "loan cap" policy in effect in a particular year. These policies assure some or all students that they will either receive no student loans in their financial aid package or receive no loans above a specified amount. Prior research has shown that such policies can result in slight increases in the enrollment of students from low-income backgrounds at private institutions (Bennett et al., 2020; Hillman, 2013), although such findings are not universal (Rosinger et al., 2018). Since there is overlap in the time periods during which institutions adopted test-optional policies and the establishment of "no-loan" and "loan cap" policies, it is valuable to control for adoption of such initiatives. The listing of "no-loan" and "loan cap" institutions and their enactment year comes from Bennett et al. (2020).

Classifying Treatment Institutions

To identify the treatment group, I proceeded through a series of steps. First, I consulted the list of test-optional selective liberal arts colleges identified in Belasco et al. (2015), which covered policies announced at selective liberal arts colleges through 2009–2010. I then reviewed the lists of test-optional institutions maintained at FairTest.org (<http://www.fairtest.org>). I examined announcements listed on FairTest.org through December 2019 and added any institutions not previously identified, as long as the institution was at least "less competitive" according to Barron's. I considered an institution as test-optional if it allowed all U.S.-based, nonhomeschooled students the choice of whether to submit test scores for consideration in the admission process, or extended the option based on requirements (e.g., GPA, class rank, intended major) that a substantial share of applicants would have met.² I

excluded institutions with test-flexible policies, which offer applicants some choice in which standardized test scores to submit (e.g., two SAT II tests rather than an SAT or ACT) but nevertheless require all students to submit some test scores for consideration during the application process.

After identifying the list of test-optional institutions that are at least somewhat selective, I determined the academic year in which the policy took effect. In addition to information gathered from reports on FairTest.org (<http://www.fairtest.org>), I consulted a list from Derousie (2014) and conducted searches in two higher education media outlets (*Inside Higher Ed* and the *Chronicle of Higher Education*) to obtain details on the policy enactment time period. I also searched regional newspapers, university websites, and school newspapers, often doing so using the Internet Archive's Wayback Machine to access archived versions of the websites. For each institution, I sought to have at least two sources to confirm the year the policy took effect (e.g., a policy first available to the class of 2018 would be linked to students entering in the 2014–2015 academic year). In 18 cases where this search process yielded only one source for the policy's timing, I contacted the institution and received clarification from an official in the admissions office (or, in one instance, library). This extensive process resulted in the list of test-optional institutions and enactment years presented in Supplemental Appendix Tables 1 and 2, available in the online version of the journal. As a result of this detailed investigation into adoption time frames, I identified several institutions where the enactment year differed from prior published research.

Empirical Strategy

The conversion to test-optional admissions is a voluntary process, with institutions adopting the policy at various points in time. Simply comparing outcome measures at test-optional institutions before and after the policy's implementation is insufficient to obtain credible estimates for the effects of test-optional policies, since unmeasured factors beyond the introduction of test-optional policies could account for part or all of the observed changes in outcomes. Several quasi-experimental methods offer approaches for addressing this challenge by assessing not just within-institution variation over time but also variation between institutions.

This study primarily relies on an econometric technique known as CITS analysis. A CITS design estimates effects by comparing average outcomes before and after an event of interest (e.g., test-optional policy enactment) for both treatment and comparison groups. Prior education research employing a CITS design has examined topics such as No Child Left Behind (e.g., Dee & Jacob, 2011; Markowitz, 2018) and school improvement grants (Hallberg et al., 2018). The CITS design can explicitly model pretreatment trends (i.e., baseline mean and baseline slope) for both treatment and comparison groups and then compare the extent to which each group varies from those

pretreatment trends following policy implementation. To model pretreatment trends, the CITS approach requires at least four time periods of data prior to policy implementation. Since the first available year of data in this study is the 2001–2002 academic year, pretreatment years extend through 2004–2005 and this analysis excludes institutions that adopted test-optional policies prior to 2005–2006.³

By explicitly modeling baseline trends, CITS differs from the more common DD design, which requires constant differences between comparison groups in the pretreatment period (as evidence of hypothetical parallel trends posttreatment in the absence of treatment). When these models are properly specified, CITS can account for even modest differences in baseline trajectories between comparison groups. In effect, one can think of the DD design as representing a special case of CITS in which the treatment and comparison group have precisely the same baseline trends. Somers et al. (2013, p. 3) have argued that the “CITS design is a more rigorous design in theory, because it implicitly controls for differences between the treatment and comparison group with respect to their baseline outcome levels and growth.”

The following model describes the main CITS design used in this study to estimate the relationship between test-optional policy implementation and outcomes:

$$Y_{it} = \beta_1 (Optional_i * Post_{it}) + \beta_2 Optional_i * Time_{it} + \alpha_t + \gamma_i + \lambda X_{it} + \epsilon_{it}. \quad (1)$$

In Equation 1, Y_{it} represents an outcome for institution i in year t , where the outcomes of interest include the number of applications, yield rate, number of Pell Grant recipients, number of FTFT URM students, and number of FTFT women enrolling. $Optional_i$ indicates whether an institution ever adopted a test-optional admissions policy during 2005–2006 to 2015–2016; $Post_{it}$ represents whether the institution was test-optional in a given year; $Time_{it}$ is centered at the year prior to test-optional policy adoption and increases one unit per year; α_t represents a year fixed effect, which is intended to account for secular trends; and γ_i indicates an institution fixed effect to represent time-invariant differences between institutions. The vector X_{it} includes a group of time-varying institutional characteristics with potential relationships to the adoption of test-optional policies as well as the outcome variables themselves. The final component of the model is the heteroskedastic-robust error term ϵ_{it} , clustered at the institution level. Based on this setup, the interaction $Optional_i * Time_{it}$ allows for different baseline linear time trends (i.e., slopes) at the earlier-adopters of test-optional policies. β_1 represents the primary coefficient of interest in this CITS model, and can be interpreted as the intercept shift between test-optional and test-requiring institutions in the years posttreatment.

Rather than solely relying on the CITS model, I also provide main estimates based on a DD design. Given the staggered timing of test-optional policy adoption, this study employs the following generalized DD model:

$$Y_{it} = \beta_1(\text{Optional}_i * \text{Post}_t) + \lambda X_{it} + \alpha_i + \gamma_t + \varepsilon_{it}. \quad (2)$$

The terms from Equation 2 largely correspond to those described in Equation 1, except that Equation 2 does not contain the centered time variable. β_1 is the primary effect estimate of interest in the DD model.

For the DD approach to yield unbiased effects, the changes in outcomes over time for the comparison group must represent what the treatment group would have experienced in the absence of the treatment. This strong presumption of equivalent changes in the absence of treatment is known as the parallel trends assumption (Angrist & Pischke, 2008). It is not possible to definitively prove or disprove that the parallel trends assumption is upheld. One method for identifying potential support for this assumption, however, is to examine outcomes in the pretreatment period, for which data are available. As shown in Supplemental Appendix Figure 1 (available in the online version of the journal), earlier- and later-adopters of test-optional policies experience broadly similar trajectories in the 4 years immediately preceding test-optional policy enactment, providing some evidence to support the parallel trends assumption. Out of caution that the parallel trends assumption may not strictly be upheld, though, and due to the largely similar results between the CITS and DD approaches, I emphasize the CITS results as my primary model.

Though the CITS and DD designs can have several distinctive elements, both fundamentally rely on timing in their estimation. Due to these methods' emphasis on policy timing, it is essential to precisely isolate the year that institutions enacted test-optional policies. One contribution of this study is its corroboration of policy enactment timing based on multiple sources, including admissions offices themselves. Even when policy timing is perfectly captured, though, it is not possible to entirely rule out effects of other changes that were coterminous with the policy of interest. For instance, institutions likely made changes to their marketing and recruitment practices in concert with test-optional policy implementation, which are therefore embedded within the estimated effects. It is also possible that test-optional policy enactment coincided with other major events, such as the introduction of a new admissions director, though the emphasis on the policy enactment year rather than policy announcement year (typically 1 or 2 years earlier) suggests a limited role for such corresponding events. Therefore, to be precise, the estimates in this study represent changes in outcomes that coincided with test-optional policy enactment, and may not necessarily isolate the causal effects solely due to test-optional policies.

Comparison Groups

Ultimately, the credibility of estimates from a CITS or DD design depends on the suitability of the comparison group used. Prior studies of test-optional policies have focused on selective liberal arts colleges (e.g., Belasco et al.,

2015), which comprised the vast majority of the earliest adopters, and therefore used test-requiring selective liberal arts colleges as a comparison group. In contrast, however, this study examines a period when a much broader pool of institutions enacted test-optional policies. For such a diverse set of institutions, no single group of institutions constitutes an obviously superior comparison group. At the same time, it is clear that the set of institutions that have voluntarily adopted test-optional policies varies from non-adopting 4-year institutions on a number of dimensions. For instance, prior research has shown that test-optional policy adoption is associated with factors such as liberal arts college status, geographic location, yield rate, and institutional selectivity (Furuta, 2017).

As one method for improving the suitability of the comparisons, the main analyses restrict the sample to a treatment group that enacted test-optional policies between 2005–2006 and 2015–2016 (“earlier-adopters”) and a comparison group that announced additional test-optional policies by December 2019 (“later-adopters”). By constraining the sample to institutions that adopted test-optional policies within this time frame, the aim is to identify a comparison group that is similar to treated institutions not just on observable dimensions but also potentially on less apparent dimensions such as institutional perceptions of merit and receptivity to change (though this can never be definitively achieved). While this restriction may aid in identifying earlier- and later adopting institutions that are highly comparable to one another, it also limits the generalizability of the findings to institutions that announced test-optional policies by December 2019, as outlined in greater detail in the Conclusion section.

In addition, this study relies on three matching procedures to identify comparison groups while accounting for factors associated with test-optional policy adoption. For the primary specification, I used a PSM procedure to identify comparison group institutions similar to the treatment group institutions on key measures available prior to policy implementation. The propensity score is a single value corresponding to the probability of an institution adopting a test-optional policy, conditional on the set of observable pretreatment covariates (Rosenbaum & Rubin, 1985). For the PSM, I used a radius matching approach, in which a treated institution’s matches include all untreated institutions with propensity scores that fall within a specified value, known as a caliper, of the treated institution’s propensity score.⁴ I used a caliper equivalent to one quarter of a standard deviation of the propensity score, as recommended in Rosenbaum and Rubin (1985). In addition to the PSM approach, I also used two other matching strategies, coarsened exact matching and Mahalanobis distance matching (Mahalanobis, 1936). For additional details on PSM and an overview of the two alternative matching methods, see the Matching Procedures section of the supplemental appendix, available in the online version of the journal.

There were five separate propensity score calculations. The first applied to all institutions in the analytic sample and matched institutions based on the

FTE number of undergraduates, tuition and fees, selectivity level, Carnegie classification, region, applications, acceptance rate, yield rate, Pell Grant recipients, FTFT URM students, and FTFT women. In addition, there were four propensity score calculations for subgroups focused on creating matches by selectivity level (less and more selective institutions) and institution type (baccalaureate college vs. master's/doctoral university). For propensity scores among the subgroups, matching variables included all the covariates used for the main PSM match, apart from selectivity, region, and Carnegie classification.

Results

This section begins with a description of the attributes of earlier- and later-adopting institutions. Afterward, I proceed with an overview of the main CITS and DD results, year-by-year (nonparametric) estimates, and robustness checks that reinforce the main findings. I conclude with several subpopulation analyses, one focused on women and the others exploring the potential for differential outcomes by selectivity level and institution type. (For reference, Tables 1–5 provide PSM-based results, while comparable estimates from coarsened exact matching and Mahalanobis distance matching approaches are available in Supplemental Appendix Tables 4–9, available in the online version of the journal.)

Descriptive Statistics

As outlined in Table 1, the set of institutions in the analytic sample that adopted test-optional policies by 2015–2016 were broadly similar to their later-adopting counterparts, although they differed on a number of observable dimensions prior to matching. Among the most conspicuous discrepancies between the two sets of institutions is the share of institutions that had a Carnegie classification as a baccalaureate liberal arts college, with 42% of the earlier test-optional institutions holding such a designation, compared to just 26% of later-adopting institutions. The matching procedure identified later-adopting institutions that more closely resembled earlier adopting institutions in terms of liberal arts designation, with a difference of 6 percentage points following PSM. Similarly, while a substantially different share of earlier and later test-optional institutions had Barron's classifications of "most competitive" or "highly competitive" in the unmatched sample (25% vs. 10%, respectively), the matched sample differed by 4 percentage points. Among the outcome measures, earlier and later test-optional institutions had notable differences in the unmatched sample on application volume (29% difference) and the acceptance rate (6 percentage points). On both measures, the matching process identified later-adopting institutions that were substantively more similar to those that implemented test-optional policies during the analytic time period. In addition to the evidence of comparison improvements due to matching shown in Table 1, Supplemental Appendix Figures 2 to 4

Table 1
Descriptive Statistics for Earlier- and Later-Adopters of Test-Optional Policies, Before and After PSM (2003)

Value	Unmatched		Matched: PSM	
	Earlier-Adopters	Later-Adopters	Earlier-Adopters	Later-Adopters
Carnegie: Baccalaureate/liberal arts	0.424 (0.497)	0.263 (0.442)	0.402 (0.493)	0.458 (0.501)
Total price (2016\$)	39,044 (5,848)	35,202 (6,160)	38,813 (5,866)	38,204 (6,496)
Barron's: most/highly competitive	0.253 (0.437)	0.102 (0.304)	0.230 (0.423)	0.189 (0.393)
Full-time equivalent undergraduates	2,389 (2,069)	2,127 (1,857)	2,370 (2,059)	2,383 (1,821)
Applications received	2,836 (2,383)	2,199 (2,519)	2,804 (2,387)	3,061 (2,849)
Acceptance rate	0.681 (0.139)	0.742 (0.144)	0.691 (0.128)	0.705 (0.143)
Yield rate	0.331 (0.100)	0.379 (0.145)	0.328 (0.101)	0.319 (0.085)
Proportion of students receiving Pell Grants	0.231 (0.128)	0.273 (0.115)	0.230 (0.123)	0.233 (0.098)
Proportion of FTFT students identified as URM	0.102 (0.121)	0.135 (0.133)	0.097 (0.115)	0.099 (0.104)
Proportion of FTFT students identified as women	0.621 (0.149)	0.627 (0.151)	0.626 (0.149)	0.622 (0.139)
Institutions (N)	99	118	90	112

Note. Earlier-adopters enacted test-optional policies between 2005–2006 and 2015–2016, while later-adopters had test-optional policies that took effect in 2016–2017 or later (and were announced by December 2019). Standard deviations are in parentheses. PSM = propensity score matching; FTFT = first-time full-time; URM = underrepresented, racially/ethnically minoritized.

Table 2
**Regression Estimates for Test-Optional Policies,
 Relative to PSM Comparison Group**

Outcome Measure	Comparative Interrupted Time Series	Difference-in-Differences
Applications (log)	0.040 (0.023) [.075]	0.031 (0.018) [.088]
Yield rate	−0.001 (0.006) [.893]	0.000 (0.005) [.971]
Pell Grant recipients (log)	0.042 (0.014) [.002]	0.031 (0.011) [.003]
FTFT URM students (log)	0.103 (0.032) [.001]	0.119 (0.026) [<.001]
FTFT women (log)	0.080 (0.014) [<.001]	0.060 (0.012) [<.001]
Institutions (<i>N</i>)	202	202

Note. Cells represent coefficient estimates from separate comparative interrupted time series and difference-in-differences models. Control variables include full-time equivalent (FTE) undergraduates (log), instructional expenditures per FTE, student services expenditures per FTE, total price, and an indicator of whether the institution had a loan-reduction initiative in effect. Standard errors in parentheses; *p* values in brackets. PSM = propensity score matching; FTFT = first-time full-time; URM = underrepresented racially/ethnically minoritized.

Table 3
**Robustness Check Results for Test-Optional Policies,
 Relative to PSM Comparison Group**

Outcome Measure	Falsification Test	Covariate Balance
Applications (log)	0.006 (0.029) [.850]	—
Yield rate	−0.006 (0.010) [.585]	—
Pell Grant recipients (log)	−0.003 (0.013) [.841]	—
FTFT URM students (log)	−0.031 (0.056) [.582]	—
FTFT women (log)	−0.003 (0.021) [.900]	—
FTE undergraduates (log)	—	0.028 (0.012) [.020]
Total price, 2016\$ (log)	—	−0.003 (0.004) [.534]
Instructional expenditures per FTE, 2016\$ (log)	—	0.017 (0.015) [.276]
Student services expenditures per FTE, 2016\$ (log)	—	0.029 (0.017) [.086]

Note. Cells represent coefficient estimates from separate comparative interrupted time series models. Standard errors in parentheses; *p* values in brackets. For the falsification test, a false test-optional adoption year is assigned. PSM = propensity score matching; FTFT = first-time full-time; URM = underrepresented racially/ethnically minoritized.

(available in the online version of the journal) illustrate that the matched sample closely aligns with the distribution of the earlier-adopters in terms of their propensity score.

Table 4
**Comparative Interrupted Time Series Regression Estimates
of Relationship Between Test-Optional Policies and Gender
Composition, Relative to PSM Comparison Group**

Outcome Measure	Comparative Interrupted Time Series
Women as a share of applicants	0.013 (0.004) [.001]
Women as a share of accepted students	0.019 (0.004) [<.001]
Women as a share of enrolling students	0.027 (0.005) [<.001]

Note. Cells represent coefficient estimates from separate comparative interrupted time series models. Standard errors in parentheses; *p*-values in brackets. PSM = propensity score matching; FTFT = first-time full-time.

Main Results

Having employed the matching procedures to identify observationally similar sets of institutions for the comparison groups, I used the CITS and DD analytic approaches to assess whether earlier-adopters of test-optional policies experienced significantly different changes in outcomes following policy enactment. Table 2 provides the resulting CITS and DD estimates for the PSM-matched institutions in the analytic sample. The first column of estimates represents results from the CITS model, with point estimates corresponding to β_1 in Equation 1. The second column provides comparable estimates from the DD model, represented by β_1 from Equation 2. As a result, the results in the first column refer to intercept shifts for earlier-adopters of test-optional policies on a particular outcome during the period when the policy was in effect, relative to later adopting institutions.

The first outcomes assessed in Table 2 are the two admissions metrics, application volume and yield rate. After controlling for time-varying characteristics and institution and year fixed effects, both the CITS and DD analyses identify modest increases of 3.1% to 4.0% in applications that are slightly above conventional levels of significance ($p = .075$ for CITS, $p = .088$ for DD). For yield rate, however, the CITS and DD models do not find evidence of a shift following test-optional policy enactment, with point estimates that are nearly zero ($p = .893$ for CITS, $p = .971$ for DD).

The second set of outcomes in Table 2 focuses on the composition of the undergraduate student body. I find indications that test-optional policy enactment is associated with increases in the enrollment of Pell Grant recipients. The detected increases for the postenactment period are relatively modest, however, amounting to 4.2% in the CITS model ($p = .002$) and 3.1% in the DD model ($p = .003$). The results in Table 2 also provide evidence that test-optional policies increased enrollment levels for URM students. These results suggest that there was a positive shift in URM enrollment following

Table 5
**Comparative Interrupted Time Series Regression Results for Test-Optional Policies by Selectivity
 Level and Institution Type, Relative to PSM Comparison Group**

Outcome Measure	Selectivity			Institution Type		
	More Selective	Less Selective	Bachelor's-Granting	Master's- or Doctorate-Granting		
Applications (log)	0.030 (0.024) [.219]	0.014 (0.044) [.757]	0.058 (0.030) [.057]	-0.019 (0.030) [.526]		
Yield rate	0.002 (0.007) [.821]	-0.002 (0.012) [.849]	-0.009 (0.009) [.345]	0.010 (0.006) [.142]		
Pell Grant recipients (log)	0.032 (0.016) [.039]	0.036 (0.022) [1.06]	0.006 (0.018) [7.42]	0.074 (0.017) [<.001]		
FTE URM students (log)	0.077 (0.035) [.031]	0.125 (0.061) [.040]	0.154 (0.045) [.001]	0.010 (0.041) [.810]		
FTE women (log)	0.046 (0.015) [.002]	0.071 (0.027) [.009]	0.081 (0.020) [<.001]	0.053 (0.018) [.003]		

NOTE: "More selective" refers to institutions listed as at least "highly competitive" according to the 2003 Barron's competitiveness index, while "less selective" refers to institutions listed as "very competitive," "competitive," or "less competitive." Cells represent coefficient estimates from separate comparative interrupted time series models. Standard errors in parentheses; *p* values in brackets. Control variables include full-time equivalent (FTE) undergraduates (log), instructional expenditures per FTE, student services expenditures per FTE, total price, and an indicator of whether the institution had a loan-reduction initiative in effect. PSM = propensity score matching; FTE = first-time full-time; URM = underrepresented racially/ethnically minoritized.

institutions' implementation of test-optional policies of 10.3% to 11.9% ($p = .001$ for CITS, $p < .001$ for DD).⁵ For the final compositional measure, estimates suggest that test-optional policies increased FTFT enrollment of women by 6.0% to 8.0% ($p < .001$ for both CITS and DD).

In addition to the relative changes for individual student groups, it is also worth considering the implications for the student body composition as a whole. As shown in Table 1, the proportion of Pell Grant recipients and URM students at matched institutions during the pretreatment period is comparatively low. As a result, the estimated increases in Pell Grant recipients and URM students amount to relatively modest gains as an absolute share of the student body. Specifically, these estimates correspond to an increase of roughly 1 percentage point in terms of both the share of students receiving a Pell Grant and the share of students from URM backgrounds. With women accounting for the majority of students at private institutions, their enrollment shift following test-optional policies represents a larger absolute increase—amounting to approximately 4 percentage points as a proportion of all FTFT students. Because the results from the CITS and DD models are quite similar, I focus on the CITS estimates from this point forward.

CITS Results by Time Since Enactment

In addition to overall estimates, I also explored potential temporal variation in the effects of test-optional policy adoption. To do so, I replaced the single $Post_t$ variable from Equation 1 with a vector of binary indicators for whether a time period was the first, second, third, fourth, or fifth or higher year of the test-optional policy (Bloom & Riccio, 2005). Figure 2 depicts the CITS results of such an estimation strategy, which focuses on the amount of time since policy enactment. While the main estimates for application volume are not measurably different from zero at conventional levels, the results in Figure 2 suggest the possibility of a short-term boost in applications that fades within a few years. For yield rate, the results appear stable and close to zero at all time points. In contrast, the point estimates for Pell Grant recipient enrollment increase over time, consistent with the fact that the measure is not restricted to first-time students. The coefficients for Pell Grant recipient enrollment are significant and positive at conventional levels beginning in the fourth year of the policy. Meanwhile, the estimates for URM enrollment are fairly consistent at each time point, although there are relatively large standard errors. The estimated effects on enrollment for women are also relatively stable, with positive point estimates across all five time periods (significant at conventional levels for three of the five).

Robustness Checks

To assess the relevance of comparison groups and the sensitivity of the estimates to different specifications, I conducted a series of robustness checks.

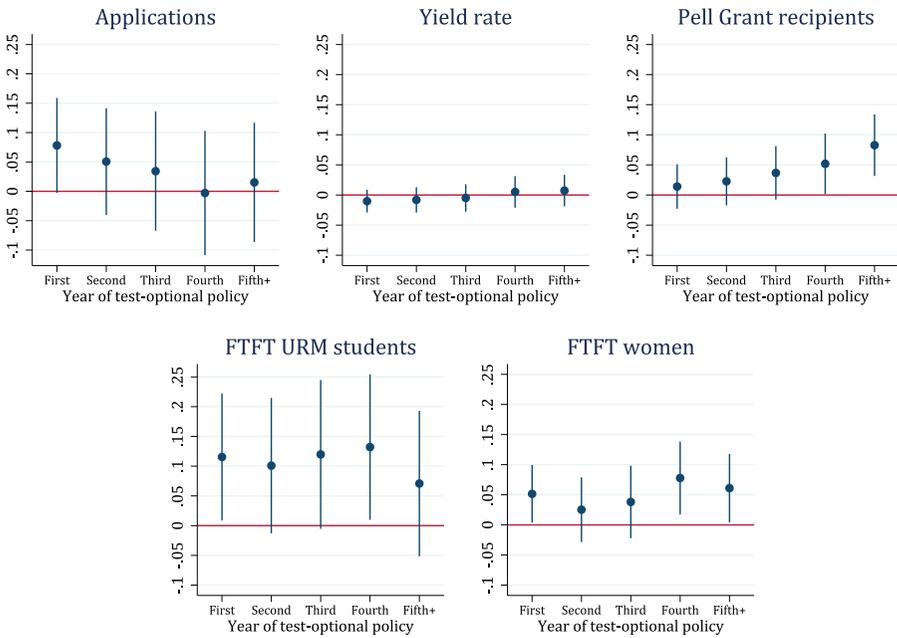


Figure 2. Year-by-year CITS estimates of relationship between test-optional adoption and key outcomes.

Note. Figures depict year-by-year CITS estimates relative to the MDM-identified comparison group. CITS = comparative interrupted time series analysis; MDM = Mahalanobis distance matching; FTFT = first-time full-time; URM = underrepresented, racially/ethnically minoritized.

The first two checks are provided in Table 3. If the CITS model isolates effects specifically attributable to test-optional policy enactment and accompanying practices in areas such as recruitment and admissions, significant effects would be discernable only after policy implementation. Effects detected in nontreatment years would suggest that any significant results identified may simply be the result of chance variation. Therefore, in the falsification test presented in the left column of Table 3, I reestimate the same CITS model but used artificial adoption years. For this falsification test, I removed data from the period of actual test-optional implementation, when it would be possible to detect actual effects. This approach does not identify any relationships based on a placebo enactment year that are significant at conventional levels. Results from the falsification test thus provide some reassurance that detected effects are unlikely to be artifacts of a spurious correlation.

As a second robustness check, the right column of Table 3 reports on covariate balance. By treating covariates as outcome measures in the CITS

model, the covariate balance check indicates whether there were significant trends in covariates that coincided with policy adoption and that therefore may be related to the main findings. Results from the covariate balance check do not suggest a clear relationship between test-optional policy enactment and the total price, instructional expenditures per FTE, or student services expenditures per FTE. There is, however, some evidence of a positive relationship between test-optional policy adoption and the FTE number of undergraduates ($p = .020$). The CITS analysis models include these measures as covariates to help account for such variation, though it is worthwhile to remain mindful of the relationship between test-optional policy adoption and this covariate.

Two additional sets of robustness checks are available and described in greater detail in the supplemental appendix, available in the online version of the journal. First, based on increasing attention to the weighted nature of DD estimates when there is variation in treatment timing (Goodman-Bacon, 2019), Supplemental Appendix Table 3 (available in the online version of the journal) displays estimates from an Oaxaca-Blinder-Kitagawa decomposition using the *bacondecomp* Stata package (Goodman-Bacon et al., 2019). This decomposition confirms that comparisons specifically between earlier- and later-adopters (i.e., the “never v. timing” group in the decomposition) account for the vast majority of the weight for the DD estimate, and that the corresponding coefficients substantially align with the main DD estimates. A second set of checks, presented in Supplemental Appendix Table 4 (available in the online version of the journal), focuses on the sensitivity of the main findings to the comparison group and CITS specification used. Across four alternative specifications in Supplemental Appendix 4 (available in the online version of the journal), I find substantially similar results to those obtained from the main CITS model.

Subpopulation Analyses

Gender

As noted earlier, rhetoric surrounding test-optional policy adoption often alludes to institutional efforts to increase representation of students from historically underrepresented backgrounds, such as Pell Grant recipients and URM students. Evidence also suggests that test-optional policy adoption may assist liberal arts colleges in improving on measures of institutional selectivity and prestige (Belasco et al., 2015), including application volume. Yet neither explanation accounts for the enrollment increases I find for women following test-optional policy adoption. Therefore, to better understand of the increase in women, I examined data on the gender composition of applicants, admitted students, and matriculants.⁶

The resulting Table 4 provides an overview of the possible steps at which test-optional policies could lead to increases in the enrollment of women.

First, test-optional policies may prompt a shift in the gender composition of the applicant pool. This change could occur if, for instance, socialization processes for women lead them to be less inclined to apply when their test scores are not in or above an institution's reported range. As shown in Table 4, there does appear to be a 1.3 percentage point increase in the share of women among all applicants following test-optional policy adoption ($p = .001$). Another possibility is that applications from women receive more favorable evaluations, on average, once test-optional policies are in effect. If, for instance, women were differentially more likely to submit applications without test scores and also had higher GPAs than men who apply, test-optional policies may contribute to higher acceptance rates for women. Consistent with this rationale, the estimates indicate that women increased an additional 0.6 percentage points at this step, resulting in a net rise of 1.9 percentage points as a share of admitted students after policy enactment ($p < .001$). Finally, women admitted under test-optional policies might enroll at different rates than men. Such differential yield rates could occur if, for instance, women were more responsive to test-optional policies as indicators of institutions' commitments to personalization. The results in Table 4 also suggest an additional gain of 0.8 percentage points for women at the enrollment stage, resulting in a cumulative increase of 2.7 percentage points as a share of enrolling students following test-optional adoption ($p < .001$). Thus, at each stage from application to acceptance to enrollment, the share of women rose, with increases in applications from women accounting for roughly half of the overall gain. These results imply the increased share of women is a net result of both factors largely under students' control (i.e., application/matriculation decisions) as well as admissions decisions.

Selectivity Level

Beyond these overall effects, it is also worthwhile to consider whether effects of test-optional policies vary by attributes of the adopting institution. In particular, since this study is among the first to include institutions from a variety of selectivity levels, it offers an opportunity to explore potential variation by institutional selectivity. Toward this end, the first two columns of Table 5 present a subgroup analysis that distinguishes between estimates for "more selective" and "less selective institutions." Here, "more selective" institutions represent the two most competitive Barron's categories, while "less selective" institutions fall within the third through fifth Barron's categories.⁷

The results of Table 5 point suggest a high degree of similarity in outcomes regardless of selectivity level. For instance, there were no consistent patterns of changes in application volume or yield rate for either group at conventional levels. In contrast, the subgroup analyses find evidence of positive intercept shifts for Pell Grant recipient enrollment following test-optional policies for both groups. These increases amount to 3.2% for the more selective

group ($p = .039$) and 3.6% for the less selective group ($p = .106$), with the latter finding more tentative. Clearer evidence emerges for enrollment of URM students and women among the more selective and less selective groups. At more selective institutions, URM student enrollment increased 7.7% following test-optional policy enactment ($p = .031$), while less selective institutions experienced increases of 12.5% for their URM student enrollment ($p = .040$). These results suggest that changes in URM enrollment patterns extended to both groups following test-optional policy enactment, with gains of a generally similar magnitude. Likewise, both more selective and less selective institutions experienced increases in women after implementing test-optional policies, with a 4.6% increase at more selective institutions ($p = .002$) and a 7.1% increase at less selective institutions ($p = .009$).

Institution Type

One of the contributions of this study is that it includes not just liberal arts colleges but also the master's and doctoral universities that adopted test-optional policies later on. Building off of this distinction, the two rightmost columns of Table 5 offer an exploration of changes by institution type, aggregated as either baccalaureate colleges or master's/doctoral universities. For baccalaureate institutions, test-optional enactment during this period was linked to increases in URM students and women, with suggestive evidence of a rise in applications. Notably, whereas prior studies did not detect gains in URM at selective liberal arts colleges using data through 2010, this more recent analysis instead finds that test-optional adoption was linked to a 15.4% increase in URM enrollment at baccalaureate institutions ($p = .001$). As in prior work, however, I find no measurable increase in Pell Grant recipients among baccalaureate institutions. Turning to the master's/doctoral universities, the subgroup analysis suggests that test-optional policies were linked to a 7.4% increase in Pell Grant recipients ($p < .001$) and a 5.3% increase in women ($p = .003$). Thus, the findings by institution type point to gains in either Pell Grant recipients or URM students, though neither baccalaureate nor master's/doctoral institutions experienced gains on both the socioeconomic and racial/ethnic diversity indicators.

Discussion

Relying on the policy adoption timing for more test-optional institutions than any prior published research, this study offers evidence on the effects of test-optional policies across the wide variety of institutions that had come to comprise the test-optional movement as of 2016. In contrast to earlier work, I find an increase of 10.3% to 11.9% in the number of URM students who matriculated following test-optional policy implementation during this era. At the same time, according to additional analyses (available upon request), there

were no detectable changes in the enrollment of White and Asian students after test-optional policies went into effect. The finding that test-optional policies increased enrollment for URM students at private institutions contributes to a broader literature on efforts to increase racial/ethnic diversity among undergraduates at selective institutions. While these increases were fairly substantial in relative terms, such effects correspond to a modest 1 percentage point increase in absolute terms in the share of URM students among the entering class. This finding suggests that test-optional policies alone may be insufficient to achieve a more transformative change in the representation of URM students at selective institutions. Such implications align with prior work showing that factors related to the higher education context, such as the proportion of same-race students at a college and the distance between home and college, are particularly salient in the college choice process for URM students (e.g., Black et al., 2020).

This study also offers some evidence supporting proponents' expectations that test-optional policies can increase socioeconomic diversity (e.g., Henson, 2014). With only a 3.1% to 4.2% increase in Pell Grant recipients, though, the effects detected are comparatively limited. Considering the baseline underrepresentation of Pell Grant recipients at the institutions examined, this shift amounts to a gain of just 1 percentage point as a proportion of all students. Given that students from low-income backgrounds are among those whose standardized test scores are systematically lower than other measures of academic performance, on average, they would appear to be some of the prime candidates to benefit from test-optional policies. Several possibilities may account for the modest change in enrollment of students from low-income backgrounds. For instance, students from low-income backgrounds may not be taking advantage of test-optional opportunities to the extent that would be beneficial, either because they are unaware of the option or because they are disposed not to use it. Such a finding would align with the conceptual framework and prior research on the college choice process of students from lower-income backgrounds, who have less access to college counseling and may have different taken-for-granted behaviors as an applicant than their higher income peers (e.g., McDonough et al., 1997). Alternatively, students from low-income backgrounds may be taking advantage of test-optional policies, but their peers strategically use test-optional policies in a manner that offsets the benefit to students from low-income backgrounds. Prior research on test score submission under test-optional policies implies the latter may be at work (Hiss & Franks, 2014), though each of these possibilities is worthy of additional investigation.

I also find that test-optional policies increased the enrollment of women, relative to matched comparison institutions. Such findings are consistent with the third channel of the proposed theory of change, with women potentially perceiving themselves as better qualified for admission under test-optional policies (even if, in fact, they were equally qualified under either policy).

Notably, due to the share of women at the private institutions examined, the absolute effects on enrollment trends for women—an increase of 4 percentage points—exceed the shifts for both Pell Grant recipients and URM students. This result sheds light on an underexamined aspect of test-optional policies, implying that test-optional policies may also prove attractive to institutions seeking to enroll larger numbers of women. Indeed, Worcester Polytechnic Institute specifically cited increasing the enrollment of women as part of its objective in launching a test-optional policy (Worcester Polytechnic Institute, 2007). With women accounting for the majority of students at adopting institutions (see Table 1), however, increases in the enrollment of women could represent an unintended consequence of the policy.

I also do not find a strong overall relationship between test-optional policy enactment and either overall application volume or yield rate. In the case of application volume, though, there is some tentative evidence of slight overall gains. These findings may reflect early gains in applications that quickly subside, as suggested in the year-by-year estimates shown in Figure 2. With dramatic escalations in the number of institutions adopting test-optional policies, however, it remains to be seen whether applicants will remain as responsive to test-optional policies in their application decisions.

Conclusion

In recent decades, a growing number and variety of institutions have turned to test-optional admissions policies. By the 2010s, what originated as a niche practice among liberal arts colleges had expanded to an increasingly mainstream approach to admissions at institutions that varied substantially on an array of attributes. These policies attracted even more extensive attention following the announcement of a test-optional policy at the University of Chicago (Anderson, 2018), one of the nation’s most selective research universities, and revelations about fraudulent standardized test scores from the Operation Varsity Blues admissions scandal (Medina et al., 2019). With large-scale test cancellations tied to the coronavirus outbreak in 2020, the shift to test-optional admissions became a practical necessity for hundreds of additional institutions (Anderson, 2020). The growing interest in and experience with test-optional admissions have made it all the more valuable to ascertain the effects of these policies.

When reflecting on the implications of this study, there are several broad points worthy of additional consideration. First, these findings depict the experiences of private institutions that enacted test-optional policies between 2005–2006 and 2015–2016, relative to others that announced policies by December 2019. Such institutions operated in an environment where students were able to decide whether to withhold their standardized test scores (with the majority still choosing to submit scores), and all of these institutions had a clear predisposition to voluntarily adopt test-optional policies. The results are unable to speak directly to the experiences of institutions that announced

test-optional policies outside the period observed, the limited number of selective public institutions that were test-optional prior to 2020, or institutions that remain test-requiring. The likely effects of test-optional policies are especially difficult to anticipate for institutions that went test-optional during the coronavirus outbreak, due in part to substantially diminished number of standardized tests completed and a rise in pass/fail grading. With dramatic reductions in test scores submitted—due to either constrained testing availability or shifting student preferences—admissions decisions place increased reliance on extracurriculars and subjective factors such as letters of recommendation. Prior research has found increased weight on such elements to be adversely related to Pell Grant recipient enrollment (Rosinger et al., 2021). Accordingly, there remains value in considering the equity implications of the admission criteria still in place at test-optional institutions and potentially expanding those criteria to include additional factors (e.g., Melguizo, 2010).

Second, the success of achieving any particular aim with a test-optional policy ultimately depends on the manner of enactment. The treatment discussed in this study is not merely the creation of a test-optional policy but also the suite of contemporaneous shifts in recruitment and admissions practices that coincided with test-optional policies. Future work on the implementation strategies at test-optional institutions may shed additional light on the mechanisms that led to the observed effects. Detailed investigation of the practices that contribute to a “successful” test-optional admissions policy may be particularly valuable for institutions that abruptly shifted to pilot test-optional policies following the coronavirus outbreak, a decision they will revisit in the years to come.

Third, modifications to standardized admissions tests may change the salience of test-optional policies. For instance, the College Board’s (2019) introduction of Landscape, which provides admissions officials with neighborhood and high school context for a student’s SAT scores, is based on an approach that has shown the potential to increase the probability of admission for students from low–socioeconomic status backgrounds (Bastedo & Bowman, 2017).

These findings suggest several avenues for future research. For example, it may be worthwhile to compare the cost and implementation burden of test-optional policies to other strategies designed to improve diversity among the student body, such as targeted outreach and recruiting or informational interventions. Institutions considering a test-optional policy (or deciding whether to retain a temporary policy) may also be interested in better understanding the mechanisms by which such increases occur. For instance, are there distinctive communication strategies at test-optional institutions that convey values that appeal to a greater number of students? Similarly, what are the specific changes in the practices of admissions officers that may help account for the observed increases in enrollment among women, URM students, and

students from lower income backgrounds? Important questions also remain about the implications of test-optional policies for outcomes beyond matriculation. For instance, if removing standardized tests from the admissions process reduces bias but adds imprecision and variability, test-optional institutions may encounter a wide variety of student needs in any particular admitted class. Therefore, a natural point of inquiry might be whether test-optional institutions are able to ensure that all admitted students have the academic supports needed to succeed. Future qualitative research would be helpful in unpacking this point. More broadly, one might consider examining whether test-optional institutions make adequate investments to develop supportive, inclusive climates for the students the policies help attract.

Overall, this study provides suggestive evidence that adopting test-optional policies can increase the enrollment of Pell Grant recipients, URM students, and women at selective private institutions, with more tenuous evidence of increases for applications. To the extent that such policies increase access for Pell Grant recipients and URM students, they help fulfill selective institutions' stated ambitions of better reflecting the socioeconomic and racial/ethnic diversity of the nation. Yet these findings also suggest that the scale of changes in demographic composition following test-optional policy adoption has been comparatively modest. For institutions seeking dramatic shifts in the student populations they serve, test-optional policies would likely need to represent one facet of a more comprehensive plan.

ORCID iD

Christopher T. Bennett  <https://orcid.org/0000-0002-8910-8098>

Notes

I thank Dominique Baker, Michael Bastedo, Jennifer Darling-Aduana, and Gary Henry for valuable feedback on earlier drafts of this article.

¹Yield rate was first available in IPEDS for 2006–2007, though I derived its value for prior years based on other available variables; I find no apparent discontinuity in yield rate values between the derived and institution-provided versions of the yield rate variable.

²I did not restrict test-optional policies to those available to all students for two reasons. First, many (perhaps most) test-optional institutions require standardized test scores for at least some applicants, such as homeschooled or international students. Second, while some test-optional institutions extend the policy only to students who meet a GPA or class rank threshold (e.g., 3.0 GPA), applicants frequently meet those requirements. Excluding institutions with GPA or class rank restrictions would overlook an important method by which institutions choose to make the test-optional policy available to applicants. The Robustness section includes a sensitivity check that excludes institutions with GPA/class rank thresholds and produces qualitatively similar estimates.

³Supplemental Appendix Table 1 (available in the online version of the journal) provides the full list of institutions excluded based on test-optional policy adoption prior to 2005–2006. The CITS requirement for 4 years of pretreatment data resulted in the exclusion of six private colleges that enacted test-optional policies between 2001–2002 and 2004–2005: Dowling, Mount Holyoke, Pitzer, Sarah Lawrence, Ursinus, and Utica.

⁴Specifically, I implemented the radius matching procedures using the `psmatch2` command in Stata Version 14.0 (Leuven & Sianesi, 2003).

⁵In additional analyses (available upon request), I also separately assess enrollment trends for individual racial/ethnic groups. Point estimates suggest increases in enrollment for Black students, Hispanic students, and Native American students alike, though large standard errors make such estimates relatively imprecise.

⁶Comparable admissions data by Pell Grant status and race/ethnicity would also be quite informative, but they are not available for the duration of the analytic time period.

⁷The results of the exploratory subgroup analysis are sensitive to the precise Barron's categories included in the "more selective" and "less selective" groups. The analyses presented limit the "more selective" group to institutions in the two most competitive Barron's categories, thereby focusing on a relatively small set of institutions traditionally regarded as having the most stringent admissions standards. For reference, these two categories included just 157 colleges and universities in the 2003 Barron's competitiveness index.

References

- Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, *49*(3), 198–211. <https://doi.org/10.3102/0013189X20902110>
- Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review*, *72*(4), 487–511. <https://doi.org/10.1177/000312240707200401>
- Anderson, N. (2018, June 14). A shake-up in elite admissions: U-Chicago drops SAT/ACT testing requirement. *Washington Post*.
- Anderson, N. (2020, October 4). Beyond "test-optional": Some "test-free" colleges drop the SAT and ACT entirely. *Washington Post*.
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Barron's Educational Series, Inc. (2002). *Barron's profiles of American colleges, 2003*.
- Bastedo, M. N., & Bowman, N. A. (2017). Improving admission of low-SES students at selective colleges: Results from an experimental simulation. *Educational Researcher*, *46*(2), 67–77. <https://doi.org/10.3102/0013189X17699373>
- Belasco, A. S., Rosinger, K. O., & Hearn, J. C. (2015). The test-optional movement at America's selective liberal arts colleges: A boon for equity or something else? *Educational Evaluation and Policy Analysis*, *37*(2), 206–223. <https://doi.org/10.3102/0162373714537350>
- Bennett, C. T., Evans, B. J., & Marsicano, C. R. (2020). Taken for granted? Effects of loan-reduction initiatives on student borrowing, admission metrics, and campus diversity. *Research in Higher Education*. <https://doi.org/10.1007/s11162-020-09615-7>
- Black, S. E., Cortes, K. E., & Lincove, J. A. (2015). Academic undermatching of high-achieving minority students: Evidence from race-neutral and holistic admissions policies. *American Economic Review*, *105*(5), 604–610. <https://doi.org/10.1257/aer.p20151114>
- Black, S. E., Cortes, K. E., & Lincove, J. A. (2020). Apply yourself: Racial and ethnic differences in college application. *Education Finance and Policy*, *15*(2), 209–240. https://doi.org/10.1162/edfp_a_00273
- Bloom, H. S., & Riccio, J. A. (2005). Using place-based random assignment and comparative interrupted time-series analysis to evaluate the jobs-plus employment program for public housing residents. *ANNALS of the American Academy of*

- Political and Social Science*, 599(1), 19–51. <https://doi.org/10.1177/0002716205274824>
- Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge University Press.
- Bowen, W. G., & Bok, D. (1998). *The shape of the river. Long-term consequences of considering race in college and university admissions*. Princeton University Press.
- Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton University Press.
- Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2018). Test scores and high school grades as predictors. In J. Buckley, L. Letukas, & B. Wildavsky (Eds.), *Measuring success: Testing, grades, and the future of college admissions* (pp. 193–226). Johns Hopkins University Press.
- Braxton, J. M. (1993). Selectivity and rigor in research universities. *Journal of Higher Education*, 64(6), 657–675. <https://doi.org/10.2307/2960017>
- Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, 163(12), 1149–1156. <https://doi.org/10.1093/aje/kwj149>
- Buchmann, C., Condrón, D. J., & Roscigno, V. J. (2010). Shadow education, American style: Test preparation, the SAT and college enrollment. *Social Forces*, 89(2), 435–461. <https://doi.org/10.1353/sof.2010.0105>
- Bulman, G. (2015). The effect of access to college assessments on enrollment and attainment. *American Economic Journal*, 7(4), 1–36. <https://doi.org/10.1257/app.20140062>
- Carmichael, L., & Mead, L. C. (Eds.). (1951). *The selection of military manpower: A symposium* (Vol. 209). National Academies Press.
- Clinedinst, M., & Koranteng, A. (2017). *2017 State of college admission*. National Association for College Admission Counseling. <http://www.nacacnet.org/global-assets/documents/publications/research/soca17final.pdf>
- College Board. (2010). *2010 college-bound senior total group profile report*. <https://research.collegeboard.org/programs/sat/data/archived/cb-seniors-2010>
- College Board. (2019). *Environmental context dashboard*. <https://professionals.collegeboard.org/environmental-context-dashboard>
- Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418–446. <https://doi.org/10.1002/pam.20586>
- Derousie, J. C. (2014). *An exploration of the diffusion and adoption of four innovations in higher education* [Doctoral dissertation]. Pennsylvania State University. <https://etda.libraries.psu.edu/catalog/23568>
- Dillon, E. W., & Smith, J. A. (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics*, 35(1), 45–66. <https://doi.org/10.1086/687523>
- Dixon-Román, E. J., Everson, H. T., & McArdle, J. J. (2013). Race, poverty and SAT scores: Modeling the influences of family income on Black and White high school students' SAT performance. *Teachers College Record*, 115(4), 1–33.
- Duncan, G. J., & Magnuson, K. A. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, 15(1), 35–54. <https://doi.org/10.1353/foc.2005.0004>
- Engle, J., & Tinto, V. (2008). *Moving beyond access: College success for low-income, first-generation students*. Pell Institute for the Study of Opportunity in Higher Education.
- Espenshade, T. J., & Chung, C. Y. (2010). *Standardized admission tests, college performance, and campus diversity*. Office of Population Research, Princeton University. <https://www.princeton.edu/~tje/files/Standardized%20AdmissionTests.pdf>

- Furuta, J. (2017). Rationalization and student/school personhood in U.S. college admissions: The rise of test-optional policies, 1987 to 2015. *Sociology of Education, 90*(3), 236–254. <https://doi.org/10.1177/0038040717713583>
- Goldfine, R. (2017, November 14). “Teen Vogue” examines value of being a test-optional college. Bowdoin College. <http://community.bowdoin.edu/news/2017/11/teen-vogue-examines-value-of-being-a-test-optional-college/>
- Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics, 98*(4), 671–684. https://doi.org/10.1162/REST_a_00600
- Goodman-Bacon, A. (2019, July). *Difference-in-differences with variation in treatment timing* [Working paper]. https://cdn.vanderbilt.edu/vu-my/wp-content/uploads/sites/2318/2019/07/24161517/ddtiming_7_24_2019.pdf
- Goodman-Bacon, A., Goldring, T., & Nichols, A. (2019). BACONDECOMP: Stata module for decomposing difference-in-differences estimation with variation in treatment timing. *Statistical Software Components*.
- Hallberg, K., Williams, R., Swanlund, A., & Eno, J. (2018). Short comparative interrupted time series using aggregate school-level data in education research. *Educational Researcher, 47*(5), 295–306.
- Hanover College. (2019). *Test-optional policy*. <https://www.hanover.edu/admission/apply/testoptional>
- Henson, M. (2014). Martha Allman: Study affirms test-optional policy for admissions. *Wake Forest Magazine*. <https://magazine.wfu.edu/2014/02/20/martha-allman-study-affirms-test-optional-policy-for-admissions/>
- Herzog, S. (2014). The propensity score analytical framework: An overview and institutional research example. *New Directions for Institutional Research, 2014*(161), 21–40. <https://doi.org/10.1002/ir.20065>
- Hillman, N. W. (2013). Economic diversity in elite higher education: Do no-loan programs impact Pell enrollments? *Journal of Higher Education, 84*(6), 806–833. <https://doi.org/10.1353/jhe.2013.0038>
- Hiss, W. C., & Franks, V. W. (2014). *Defining promise: Optional standardized testing policies in American college and university admissions*. National Association for College Admission Counseling.
- Hoxby, C. M., & Avery, C. (2012). *The missing “one-offs”: The hidden supply of high-achieving, low income students* (NBER Working Paper No. 18586). National Bureau of Economic Research.
- Hoxby, C. M., & Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review, 105*(5), 514–517. <https://doi.org/10.1257/aer.p20151027>
- Hoxby, C. M., & Turner, S. (2019). *Measuring opportunity in US higher education* (NBER Working Paper No. 25479). National Bureau of Economic Research.
- Hurlburt, S., Peek, A., & Sun, J. (2017). *Delta Cost Project database 1987–2015*.
- Hyman, J. (2017). ACT for all: The effect of mandatory college entrance exams on post-secondary attainment and choice. *Education Finance and Policy, 12*(3), 281–311. https://doi.org/10.1162/EDFP_a_00206
- Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association, 106*(493), 345–361. <https://doi.org/10.1198/jasa.2011.tm09599>
- Karavel, J. (2006). *The chosen: The hidden history of admission and exclusion at Harvard, Yale, and Princeton*. Houghton Mifflin Harcourt.
- Klasik, D. (2012). The college application gauntlet: A systematic analysis of the steps to four-year college enrollment. *Research in Higher Education, 53*(5), 506–549. <https://doi.org/10.1007/s11162-011-9242-3>

- Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. Farrar, Straus & Giroux.
- Leonard, D. K., & Jiang, J. (1999). Gender bias and the college predictions of the SATs: A cry of despair. *Research in Higher Education, 40*(4), 375–407. <https://doi.org/10.1023/A:1018759308259>
- Leuven, E., & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. <http://ideas.repec.org/c/boc/bocode/s432001.html>
- Light, A., & Strayer, W. (2000). Determinants of college completion: School quality or student ability? *Journal of Human Resources, 35*(2), 299–332. <https://doi.org/10.2307/146327>
- Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India, 2*(1), 49–55.
- Manski, C. F., & Wise, D. A. (1983). *College choice in America*. Harvard University Press.
- Markowitz, A. J. (2018). Changes in school engagement as a function of No Child Left Behind: A comparative interrupted time series analysis. *American Educational Research Journal, 55*(4), 721–760. <https://doi.org/10.3102/0002831218755668>
- Mattern, K. D., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). *Differential validity and prediction of the SAT®* (Research Report No. 2008-4). College Board.
- McDonough, P. M., Korn, J. S., & Yamasaki, E. (1997). Access, equity, and the privatization of college counseling. *Review of Higher Education, 20*(3), 297–317. <https://doi.org/10.1353/rhe.1997.0004>
- Medina, J., Benner, K., & Taylor, K. (2019, March 12). Actresses, business leaders and other wealthy parents charged in U.S. college entry fraud. *New York Times*.
- Melguizo, T. (2010). Are students of color more likely to graduate from college if they attend more selective institutions? Evidence from a cohort of recipients and non-recipients of the Gates Millennium Scholarship Program. *Educational Evaluation and Policy Analysis, 32*(2), 230–248. <https://doi.org/10.3102/0162373710367681>
- Mohr, T. S. (2014, August 25). Why women don't apply for jobs unless they're 100% qualified. *Harvard Business Review*. <https://hbr.org/2014/08/why-women-dont-apply-for-jobs-unless-theyre-100-qualified>
- Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives, 24*(2), 129–44. <https://doi.org/10.1257/jep.24.2.129>
- Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics, 33*(2), 493–520. <https://doi.org/10.1086/678520>
- Perna, L. W. (2006). Studying college choice: A proposed conceptual model. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 21, pp. 99–157). Springer.
- Plank, S. B., & Jordan, W. J. (2001). Effects of information, guidance, and actions on postsecondary destinations: A study of talent loss. *American Educational Research Journal, 38*(4), 947–979. <https://doi.org/10.3102/00028312038004947>
- Reardon, S. F., Kalogrides, D., & Shores, K. (2019). The geography of racial/ethnic test score gaps. *American Journal of Sociology, 124*(4), 1164–1221. <https://doi.org/10.1086/700678>
- Robinson, K. J., & Roksa, J. (2016). Counselors, information, and high school college-going culture: Inequalities in the college application process. *Research in Higher Education, 57*(7), 845–868. <https://doi.org/10.1007/s11162-016-9406-2>

- Robinson, M., & Monks, J. (2005). Making SAT scores optional in selective college admissions: A case study. *Economics of Education Review*, *24*(4), 393–405. <https://doi.org/10.1016/j.econedurev.2004.06.006>
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38. <https://doi.org/10.1080/00031305.1985.10479383>
- Rosinger, K., & Ford, K. S. (2019). Pell Grant versus income data in postsecondary research. *Educational Researcher*, *48*(5), 309–315. <https://doi.org/10.3102/0013189X19852102>
- Rosinger, K. O., Belasco, A. S., & Hearn, J. C. (2018). A boost for the middle class: An evaluation of no-loan policies and elite private college enrollment. *Journal of Higher Education*, *90*(1), 27–55. <https://doi.org/10.1080/00221546.2018.1484222>
- Rosinger, K. O., Sarita Ford, K., & Choi, J. (2021). The role of selective college admissions criteria in interrupting or reproducing racial and economic inequities. *The Journal of Higher Education*, *92*(1), 31–55. <https://doi.org/10.1080/00221546.2020.1795504>
- Rosner, J. (2012). The SAT: Quantifying the unfairness behind the bubbles. In J. Soares (Ed.), *SAT wars: The case for test-optional college admissions* (pp. 104–117). Teachers College Press.
- Rubin, D. B., & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, *52*(1), 249–264. <https://doi.org/10.2307/2533160>
- Saboe, M., & Terrizzi, S. (2019). SAT optional policies: Do they influence graduate quality, selectivity or diversity? *Economics Letters*, *174*, 13–17. <https://doi.org/10.1016/j.econlet.2018.10.017>
- Somers, M. A., Zhu, P., Jacob, R., & Bloom, H. (2013). *The validity and precision of the comparative interrupted time series design and the difference-in-difference design in educational evaluation*. MDRC. https://www.mdrc.org/sites/default/files/validity_precision_comparative_interrupted_time_series_design.pdf
- Sweitzer, K., Blalock, A. E., & Sharma, D. B. (2018). The effect of going test-optional on diversity and admissions. In J. Buckley, L. Letukas, & B. Wildavsky (Eds.), *Measuring success: Testing, grades, and the future of college admissions* (pp. 288–307). Johns Hopkins University Press.
- Syverson, S. (2007). The role of standardized tests in college admissions: Test-optional admissions. *New Directions for Student Services*, *2007*(118), 55–70. <https://doi.org/10.1002/ss.241>
- Syverson, S., Franks, V., & Hiss, W. (2018). *Defining access: How test-optional works*. National Association for College Admission Counseling. <https://www.nacacnet.org/globalassets/documents/publications/research/defining-access-report-2018.pdf>
- Tebbs, J., & Turner, S. (2005). College education for low-income students: A caution about using data on Pell Grant recipients. *Change: Magazine of Higher Learning*, *37*(4), 34–43. <https://doi.org/10.3200/CHNG.37.4.34-43>
- U.S. Department of Education. (2018). *Distribution of Federal Pell Grant Program funds*. <https://www2.ed.gov/finaid/prof/resources/data/pell-institution.html>
- Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*(9), 1132–1139. <https://doi.org/10.1111/j.1467-9280.2009.02417.x>
- Warikoo, N. K. (2016). *The diversity bargain: And other dilemmas of race, admissions, and meritocracy at elite universities*. University of Chicago Press.

- Wechsler, H. S. (1977). *The qualified student: A history of selective college admission in America*. Wiley.
- Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT® scores, high school grades, and SES. *Educational Assessment, 20*(1), 23–45. <https://doi.org/10.1080/10627197.2015.997614>
- Worcester Polytechnic Institute. (2007, May 8). *Worcester Polytechnic Institute is first nationally-ranked science and engineering university to make SAT optional for admissions*. <https://www.wpi.edu/news/satoptional>
- Zernike, K. (2015, October 24). Obama Administration calls for limits on testing in schools. *New York Times*. <https://www.nytimes.com/2015/10/25/us/obama-administration-calls-for-limits-on-testing-in-schools.html>
- Zwick, R. (Ed.). (2004). *Rethinking the SAT: The future of standardized testing in university admissions*. Routledge.
- Zwick, R. (2019). Assessment in American higher education: The role of admissions tests. *ANNALS of the American Academy of Political and Social Science, 683*(1), 130–148. <https://doi.org/10.1177/0002716219843469>

Manuscript received August 22, 2019

Final revision received February 17, 2021

Accepted February 22, 2021