Spring 5-1-2021

The Growth of the Test-Optional Movement: Analysis of Test-Optional Admissions Policies in American Higher Education

Angela C. Lofaro
angela.lofaro@shu.edu

The Growth of the Test-Optional Movement:
Analysis of Test-Optional Admissions Policies in American Higher Education

Angela C. Lofaro

Dissertation Committee

Dr. Eunyoung Kim, Mentor
Dr. Richard Blissett, Committee Member
Dr. Alyssa McCloud, Committee Member

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Education Leadership, Management, and Policy
Seton Hall University

2021

COLLEGE OF EDUCATION & HUMAN SERVICES
DEPARTMENT OF EDUCATION LEADERSHIP MANAGEMENT & POLICY

APPROVAL FOR SUCCESSFUL DEFENSE

Angela C. Lofaro has successfully defended and made the required modifications to the text of the doctoral dissertation for the **Ph.D.** during this **Fall** Semester **2020**.

DISSERTATION COMMITTEE
(please sign and date)

Dr. Eunyoung Kim
Mentor **Date**

Dr. Richard Blissett
Committee Member **Date**

Dr. Alyssa McCloud
Committee Member **Date**

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation.

Abstract

The landscape of American higher education is becoming more competitive each year. Discussions about equity and increasing access to underrepresented student groups grows each year, particularly regarding the role of standardized testing. Some research has shown the SAT and ACT exams, the two primary college entrance exams used in the United States, to be discriminatory towards students of some racial and ethnic minority and low-socioeconomic backgrounds. The number of institutions in the U.S. choosing to forego standardized test scores as part of their admission process for freshman applicants is growing annually, particularly in the midst of the global COVID-19 pandemic beginning in 2020.

This study sought to identify whether or not adopting a test-optional freshman admission policy was associated with changes in specific institutional outcomes. The research questions focused on outcomes in the areas of diversity in enrollment, reputation of the institution, and student success. Data were obtained from the National Center for Education Statistics' Integrated Postsecondary Educational Data System (IPEDS) as well as *U.S. News and World Report* and were analyzed using fixed-effects panel regression models. The sample comprised 1,681 4-year, degree-granting institutions in the United States, and the longitudinal dataset spanned 8 years.

The results of the analysis indicated statistically significant results for five of the six dependent variables when institutions switched from a test-required to a test-optional admission policy when controlling for other predictors included in the models; only the *U.S. News and World Report* rank variable was not found to be significantly related to test score policy.

Keywords: test-optional, admissions, college access, diversity, enrollment

Acknowledgements

I want to sincerely thank Dr. Eunyoung Kim, my dissertation mentor, who has patiently guided me on this journey from start to finish. Even after she left the faculty of Seton Hall University, she continued to devote her own time to mentoring me to the conclusion of my degree. I cannot thank her enough for her years of commitment to this project and for her excellent notes, which always pushed me to become a better writer.

I would like to thank Dr. Richard Blissett, a member of my committee, who aside from providing feedback and notes went particularly above and beyond in assisting me with my statistical analyses. He willingly jumped in to help when computer software failed me, and I could not have concluded my analysis without his great help.

I want to thank Dr. Alyssa McCloud, a member of my committee and my professional mentor. Dr. McCloud is the person who most encouraged me to pursue a doctoral degree in higher education, and it is because of her advice that after much stalling I finally decided to apply to the program. In addition, she is the person who taught me that higher education administration is much more than just a job but rather a true career path and a calling. I greatly appreciate her approach to supervision, which includes much teaching and professional development, and it is mostly to her that I owe my decision to adopt higher education administration, and admissions in particular, for my own career.

I would also like to thank my parents, Joe and Mimi Berry, who have always believed in me and accepted without question that I could do anything I set out to do. They have been tireless examples to me of what hard work can achieve, both professionally and in one's personal life. I thank them both for giving me such a supportive environment in which to grow up and thrive, and for continuing to guide me even in adulthood.

My aunt, Dr. Nancy Durbin, has also been an inspiration to me. She was the first person I knew who earned a doctoral degree, and I remember visiting her office as a child to see her new name plaque on the door sporting the title of "Dr." for the first time. I thank her for being one of my first examples of an empowered, educated, confident woman in a professional field. Growing up with an aunt who had earned a doctoral degree in a challenging subject made me believe that it was totally normal for one to do such a thing, and taught me that anything was possible.

Finally, I cannot end my acknowledgements without giving the greatest thanks of all to my husband, Arcangelo Lofaro. He has supported me through the good times, pushed me through the hard times, and never allowed me to give up even when I wanted to do so. He knows when to be tough on me and when to allow a little leniency, and throughout this entire experience he never once believed I would not make it to the end. I cannot thank him enough for walking this road alongside me. This achievement would not be possible without him.

Dedication

To my children, Giada and Bruno, for whom I strive every day to be the best version of myself that I can be: May you forever have a love of learning, and may you always know the value of hard work and perseverance. You can do anything you put your minds to and your father and I will be here to support you all the way.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction and Background

Introduction

In recent years the landscape of the American higher education system has found itself under great scrutiny by students, policymakers, the media, and the general public. With the rising costs of college education, institutions are increasingly pressured to justify student outcomes and deal with the increasingly competitive nature of college admissions. Coupled with these greater issues lies a central theme that resurfaces in the debate repeatedly: *access to college*.

Providing broader access to college has been the central focus of a number of empirical studies. Scholars identify many barriers to college that stand between potential students and the dream of higher education. They have studied college access by race and ethnicity (e.g., Buchmann et al., 2010; Card & Rothstein, 2007; Kobrin et al., 2007; Perez, 2002; Walpole et al., 2005), gender (e.g., Buchmann et al., 2010; Kobrin et al., 2007; Micceri, 2010; Perez, 2002), socioeconomic status (e.g., Buchmann et al., 2010; Geiser & Santelices, 2007; Kobrin et al., 2007), and more. Many in the education field believe that standardized tests might pose an additional barrier to college for students from disadvantaged backgrounds.

History of Standardized College Entrance Exams: SAT and ACT

Many countries around the world have codified a standardized secondary school system including a national exam to be completed by all students prior to graduating high school. However, the United States enjoys no such standardization. Instead, education at all levels from kindergarten through college is overseen by individual states rather than the federal government. As such, the U.S. has no national exam that all high school students must pass in order to graduate from secondary school.

In the absence of a nationalized system of testing high school graduates, two primary college entrance exams have been developed independently during the past century to fill this void: the SAT and the ACT. The parent organization of each exam claims that its test is meant to provide a standardized way of evaluating college applicants on common ground (Buchmann et al., 2010). A student's test scores are supposed to give college admission officers insight into the level of each applicant's preparedness for success in college. Yet while a college education continues to be viewed as the best path to upward social and economic mobility in the United States, underrepresented minority groups still face obstacles in the admission process.

The older of the two standardized tests is the SAT, previously known as the Scholastic Aptitude Test or the Scholastic Achievement Test (Rosner, 2012, p. 104), now simply named by its acronym. The SAT is administered by the College Board (Applebome, 1997), a nonprofit organization founded in 1900 to standardize the use of college entrance examinations used in the United States (Zwick, 2007a). Though the College Board itself has been in existence since 1900, the first iteration of the SAT exam originated in 1926 (Atkinson & Geiser, 2012, p. 26). The test was originally created to provide colleges with a means of evaluating the raw intellect of their applicants rather than the extent to which they had succeeded in learning their high school material (Kaplan, Inc., 2016). Harvard University became the first American higher education institution to use SAT scores as part of its admission process in 1934 under the leadership of then-president James Bryant Conant, and the trend spread to other institutions throughout the 1930s, 1940s, and beyond. (Kaplan, Inc., 2016). By 1957 over 50% of college-bound students took the SAT annually, by the 1990s the exam had become pervasive, and in the 21st century more than two million high school students take the SAT exam each year (Buchmann et al., 2010).

In recent decades both the format and scoring of the SAT have been updated several times, most notably in 1994, 2005, and 2016 (Kaplan, Inc., 2016). The current version of the exam, which was first administered in March 2016, consists of two major sections: Math and Evidence-Based Reading and Writing, with a scale of 200–800 points in each section. An optional essay section is also available for students to complete, which colleges may or may not include in their application review processes (College Board, 2020c).

For years, the SAT remained the primary college entrance exam in the United States, but in 1959 a rival test emerged on the scene: the ACT (Atkinson & Geiser, 2012, p. 30). Developed by Professor E. F. Lindquist of the University of Iowa, the ACT exam was built upon an entirely different philosophy from the SAT. Originating not only as a means of providing competition for the SAT, the ACT was also intended to establish an assessment for how well students had learned what was being taught in their high schools rather than their innate aptitude (Atkinson & Geiser, 2012, p. 30; Perez, 2002). The ACT exam is owned by a parent company of the same name as the test: ACT, Inc. The test comprises four primary sections: English, mathematics, reading, and science, along with an optional writing section (ACT, Inc., 2020b). Scores are calculated on a scale of 1–36 for each of the first four sections, the average of which is reported as the overall "composite" score for the student (ACT, Inc., 2020b). Notably, the ACT includes a section on science within its exam which the SAT does not (ACT, Inc., 2020b).

Although the ACT and SAT evaluate students on completely different platforms, pose different questions, cover some separate academic subjects, and score on dissimilar scales, most postsecondary institutions in the United States that require applicants to submit some type of standardized test scores will accept results from either or both exams. Concordance tables equating scores from one test to the other can be found on both organizations' websites, although

the most recent concordance table for the ACT to the 2016 revised SAT was evidently developed independently by the College Board without the input of the ACT organization (Jaschik, 2016), leaving its validity in question. Historically, the SAT has been the far more popular test taken by students along each of the coasts and throughout parts of the South, whereas the ACT has dominated the testing landscape in the Midwest (Thomas, 2004). However, since 2003 this trend has shifted as more students throughout the country opt to take both exams in the hopes of achieving better scores on one test or the other, thereby increasing the chance of being admitted to a more selective college or university (Thomas, 2004). In 2012, both the College Board and ACT, Inc. recorded more than 1.6 million college-bound students taking their respective exams, with many students overlapping and taking both exams at least once (Lewin, 2013).

Although there is much debate about the value in doing so, average SAT and/or ACT scores for public secondary schools are often reported at the school, district, or state level. In many cases this measurement is used as one way of evaluating the quality of the educational institution in question (Grissmer, 2000). Because they are utilized in such a way, the average test scores of an individual high school or entire school district are sometimes applied to calculations establishing funding levels of public schools and districts (Grissmer, 2000). Although other estimations of the K–12 educational system in the United States indicate that the quality has improved over the past 25 years particularly for minority students, average test scores have decreased over this period (Grissmer, 2000), which is another reason why many critics feel that the exams are inherently flawed in many ways since they do not appear to accurately assess the current state of secondary education in the United States. This practice of using scores to assess funding levels creates what some believe to be an inequitable method of funding, which relies too heavily upon the wrong metrics to determine allocations (Grissmer, 2000).

Problem Statement

Regardless of the intentions behind the exams, empirical research has shown that these standardized tests are discriminatory in several different ways, most notably against minority, low-income, and female students who all tend to score lower than their counterparts on both the ACT and SAT (Mattern et. al., 2011; Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Rosner, 2012, p. 104; Syverson, 2007; Zwick, 2007a; Zwick, 2007b). Fewer students from lower income brackets have the opportunity to attend college for a variety of reasons and, when they do enroll, these more vulnerable populations are more likely to attend less selective schools (Buchmann et al., 2010). Every year between 2000 and 2015, more students from high-income families attended college immediately following high school graduation than students from low- and middle-income families (National Center for Education Statistics [NCES], 2017). In fact, in 2015 eighty-three percent of students from high-income families went on to immediately enroll in college, whereas just 63% of students from low- and middle-income families did so (National Center for Education Statistics, 2017).

In response to this criticism, the testing organizations themselves have published research attempting to prove the validity of their exams and support the mainstream belief that standardized test scores are a necessary element of the American college admission process (Atkinson & Geiser, 2012, p. 39; Kobrin et al., 2007; Kobrin et al., 2008). The fact that each agency has published work overwhelmingly supporting the validity of its own exam is unsurprising but raises the question of conflict of interest (Atkinson & Geiser, 2012, p. 39). Despite these efforts to maintain their stake in the college admission industry, hundreds of colleges and universities across the United States have opted to de-emphasize the importance of SAT/ACT scores in their application review policies (Schaeffer, 2012, p. 153). Some institutions

have achieved this by making test scores optional for some of their applicants, whereas other institutions have removed this element from their application requirements altogether by refusing to consider standardized test scores at all (Schaeffer, 2012, p. 165). Although the moniker for this movement as a whole is "test-optional," the term actually describes the full range of policies schools have adopted from making standardized test scores non-compulsory for just a portion of applicants to removing them from the admission process entirely (Syverson, 2007).

The National Center for Fair and Open Testing (FairTest) is a nonprofit organization that devotes its efforts to advocating for fair testing policies as they relate to education in the United States. FairTest believes that there are many flaws inherent in both the concept of standardized testing as well as the specific standardized tests used in the American educational system such as the SAT, ACT, and GRE, for example. Their overarching goal is to increase access to education by removing the barrier that standardized testing creates. The organization puts "…special emphasis on eliminating the racial, class, gender, and cultural barriers to equal opportunity posed by standardized tests, and preventing their damage to the quality of education" (National Center for Fair and Open Testing [FairTest], 2018).

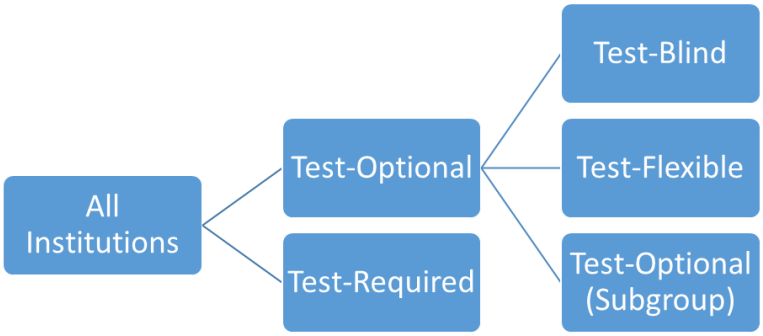
Every higher education institution can be labeled as either test-optional or test-required. From there, FairTest groups test-optional schools into three different categories based on their standardized testing policy: test-flexible, test-blind, and test-optional. Figure 1 illustrates this variety of test score policies. According to Bob Schaeffer of FairTest, Hampshire University is the only institution that is considered to be truly test-blind,[1] meaning it does not use standardized

[1] Sarah Lawrence College also had a test-blind admission policy for a number of years before returning to a test-optional policy in 2012, giving applicants the choice of whether or not to submit SAT and/or ACT scores with their freshman applications (Jaschik, 2012).

test scores of any kind to evaluate its applicants (Stewart, 2016). Espenshade and Chung (2012) refer to the same policy as "Don't Ask, Don't Tell" (p. 187).

Figure 1

Categories of College Admissions Testing Policies



Approximately 10 schools in the list are test-flexible (Stewart, 2016). Test-flexible institutions require some type of standardized test but it does not have to be the SAT or ACT. For example, some institutions will accept the SAT II subject test, Advanced Placement scores, or International Baccalaureate scores in lieu of the typical SAT or ACT exam (Dixon, 2015). The key difference here is that students must still submit some type of score, but they have flexibility in determining which test(s) they wish to include.

The remaining group of colleges and universities that are neither test-blind nor test-flexible fall into the subcategory of test-optional schools, which is unfortunately labeled with the same name as the larger group. These are the institutions that do not require ACT or SAT scores for every single applicant, but they differ in the form that their test-optional policies take. Some schools request an additional essay or writing sample to substitute for test scores. Others request graded papers or projects from high school to be submitted with the application. Still others make interviews mandatory thereby putting more emphasis on identifying the students for whom

the school is the best fit on a more personal level rather than strictly by looking at quantitative metrics on paper (see Table 1 for more details). In some cases, the institution imposes the alternate requirements upon most or all of their applicant pools, whereas others may only make test scores optional for students who meet a certain minimum high school GPA or some other exempting benchmark (Allman, 2012, p. 169).

Table 1

Summary of Alternative Testing Policies

Policy	**Definition**	**Notes**
Test-Optional (Subgroup)	Institution does not require SAT or ACT scores for some if not all of its applicants; for students to whom this option is available, it is their choice whether or not to submit scores	May be optional for everyone or may be optional just for students who meet certain benchmarks (e.g., a minimum HSGPA requirement); may have other requirements in lieu of SAT/ACT scores such as additional essays, graded high school assignments, or mandatory interviews.
Test-Flexible	Some if not all applicants have the option to submit alternative test scores in lieu of SAT/ACT scores	E.g., SAT II, Advanced Placement, International Baccalaureate, or TOEFL scores, or some combination of these
Test-Blind (aka "Don't Ask, Don't Tell"	Institution reviews no standardized test scores for any applicant	Even if scores are submitted, they are not taken into consideration during the applicant review process; may not apply to non-native English speakers who must still submit English language tests (e.g., TOEFL, IELTS).

Note. Information adapted from Allman (2012, pp. 169), Dixon (2015), Espenshade & Chung (2012), Stewart (2016), Syverson (2007), and the National Center for Fair and Open Testing [FairTest] (2020).

On its website, FairTest maintains a running list of all the institutions of higher education within the United States that currently participate in test-optional admissions. The list contains "…bachelor degree granting institutions that will not require recent U.S. high school graduates to submit ACT or SAT results to be eligible for fall…admission" (National Center for Fair and Open Testing [FairTest], 2020), and as of fall 2020 the list contained more than 1300 institutions. The colleges and universities within said list fall into a wide variety of categories and types as defined by the *U.S. News and World Report* rankings, which uses Carnegie Classifications to categorize institutions in order to rank them against their compatible peer institutions. FairTest's list includes more than 925 test-optional institutions that are ranked highly within their respective *U.S. News and World Report* categories (National Center for Fair and Open Testing [FairTest], 2020).

Although the belief that the SAT and ACT exams are discriminatory is a major reason for adopting a test-optional admission policy (Robinson & Monks, 2005; Syverson, 2007), colleges and universities choose this route for a variety of reasons ranging from ethical to tactical. Schools may look at a test-optional policy as a vehicle for achieving other objectives such as enrollment, diversity, or academic quality goals. The question is whether or not test-optional admission policies indeed assist an institution in achieving any or all of these objectives, and moreover, is the test-optional movement contributing towards increased access to higher education for racial minorities and other underrepresented populations?

Although a handful of institutions' test-optional policies date back to the 1960s, the trend has experienced a significant increase since the beginning of the 21st century (Morgan, 2016). Several colleges have conducted their own analysis on the effects of such changes to their admission procedures in order to justify either the program's continuance or stoppage. Case

studies conducted either by the institutions themselves or by outside scholars dominate the limited research available on the outcomes of test-optional admission policies with only a few large-scale empirical studies included (Espenshade & Chung, 2012, pp. 185–86), the most notable of these being Espenshade and Chung in 2012 Belasco et al. in 2015, and Syverson et al. in 2018.

Belasco et al. (2015) examined the impact of test-optional policies on diversity rates at liberal arts institutions from 1992–2000 using a narrow definition of test-optional; they only categorized those institutions as test-optional, which had made test score submission optional for *all* applicants, narrowing their sample to exclude those schools that have made submission optional for just a portion of their applicant pools even if it was an option for the majority (Belasco et al., 2015). The results of this study demonstrated that the proportion of minority students and low-SES students enrolled, as measured by Pell Grant recipients, did not increase among the 180 schools in the sample. More significantly, these two groups did not overtake these diversity rate indicators at test-requiring schools, but rather remained less than the rates at test-requiring institutions (Belasco et al., 2015). Test-optional institutions did experience increases in total application numbers and average SAT scores over the years of the study, but it should be noted that these measures increased at test-requiring institutions, as well, albeit by lesser rates (Belasco et al., 2015).

Espenshade and Chung (2012) studied diversity rate outcomes of test-optional policies in public and private institutions using data from the National Study of College Experience (NSCE) database. The study involved running simulations with the NSCE data on the public and private schools separately. In contrast to the Belasco et al. (2015) study, Espenshade and Chung found that the diversity of institutions as defined by underrepresented minorities and low-SES students

did increase for both sets of institutions when adopting a test-optional policy but more so at private institutions than public institutions. The primary difference between the executions of the two aforementioned studies is the level analysis as Belasco and his colleagues scrutinized institutional-level data, whereas Espenshade and Chung utilized student-level data in their study.

The 2018 study by Syverson et al. took a different approach to analyzing the outcomes of test-optional admissions. The authors collected student-level data from 28 institutions of varying size for the 2 years prior to adopting a test-optional policy and the 2 years after the policy was enacted, with a few exceptions due to availability of data. This study not only examined some general trends within the entire dataset itself, but its primary focus was to match each individual test-optional school in their sample to a specific test-requiring peer institution with similar characteristics in order to do a direct comparison of the two test score policies. Their principal findings as they pertain to the research questions of this study were that, compared to individual test-requiring peers, test-optional institutions experienced an increase in applications, an increase in underrepresented minority students applying and enrolling, and a small increase in Pell Grant recipients applying and enrolling (Syverson et al., 2018).

Additionally, Hillary Morgan's (2016) dissertation examined the interaction between submitting standardized test scores or not and student sensitivity to financial aid in enrollment at a single Northeastern liberal arts college. Her study found that students were more likely to enroll at the sample institution when offered institutional aid, regardless of whether it was more merit- or need-based aid. Interestingly, non-submitters were more likely to enroll with an extra $1000 of aid offered than test score submitters (Morgan, 2016).

Each of these previous studies, along with my additional research, has informed my knowledge of the test-optional movement in a different way. One study mentioned above looked

at a very narrow definition of being test-optional, another looked at diversity outcomes using student-level data, the third used a peer-matching technique to directly compare outcomes of test-score policies, and the last integrated financial aid into the equation while focusing on a single institution. In this study, I seek to combine elements of each of these to create a large-scale empirical study that has not been performed to date: looking at diversity outcomes in addition to student success outcomes using a broader definition of test-optional; any institution that allows at least a portion of their applicants to apply without submitting SAT or ACT scores is categorized as test-optional in this study. In an effort to meaningfully contribute to the ongoing research and understanding of the effects of the test-optional movement on the American higher education system, this quantitative study seeks to dive deeper into the topic of test-optional admission policies guided by the following questions:

1. Are the diversity rates of test-optional institutions significantly different from test-requiring institutions, as measured by the proportion of racial and ethnic minority students and Pell Grant recipients enrolled?

2. Do institutions with test-optional policies gain a competitive advantage in reputation and selectivity as measured by a decrease in acceptance rates and/or an increase in national rankings in the *U.S. News and World Report*?

3. What impact, if any, does the adoption of a test-optional admission policy have on student success as measured by first-year retention and 6-year graduation rates?

Significance of the Study

As this dissertation is being finalized in 2020, the world is experiencing the unprecedented, global COVID-19 pandemic. Because many SAT and ACT exam dates have been cancelled this year, hundreds of colleges and universities in the United States have opted for test-optional admission policies for at least fall 2021 first-year applicants. This has resulted in

a rapid increase in the discussion of what test-optional policies mean and how to implement them on various campuses in a very short timeframe.

The information publicly available on test-optional admissions heavily favors case studies and reports of outcomes at the individual institutional level that have become test-optional, and very little empirical research exists on the large-scale and long-term impacts of adopting such a policy throughout a large proportion of American colleges and universities. Many institutions seem to adopt a test-optional policy based upon simple analysis of admission statistics for a given institution, anecdotal evidence from peer institutions, or a leap of faith. Institutions often track the results of their own test-optional experiment for a few years and base their decision to continue it or not on the results of that analysis. However, if more large-scale, evidence-based studies were produced on the concrete outcomes of the test-optional movement, particularly in the areas of diversity and student success, around which much current policy discourse in higher education has been centered in recent years, institutions may be more well informed before deciding to become test-optional or not.

Controversy over the SAT and ACT will be discussed at length in the following chapter, particularly in reference to their bias against racial and ethnic minority students. If indeed the tests are discriminatory, and significant results that support the idea of test-optional admissions can be found by studies such as this one, the case for increasing minority students' access to college by eliminating the barrier of standardized entrance exams becomes even stronger.

Organization of the Dissertation

In the following chapter of this dissertation, I review the existing literature on test-optional admissions and the policy's outcomes, as well as related topics to the present discussion including the ability of the SAT and ACT to predict college success (or not), the argument that

the SAT and ACT are discriminatory, and the role the testing industry and test preparation play in the overall picture. In Chapter III, I discuss my methodology along with a description of the variables used in my analysis, a summary of the statistical tests employed to reach my results, and a discussion of possible limitations of the study. Chapter IV presents an analysis of the results of the study followed by the concluding chapter that includes a discussion of those results along with suggestions for future research on this topic.

Chapter II

Literature Review

Introduction

The discussion of test-optional admission policies continues in this chapter with a review of the existing literature on the subject. As mentioned before, only a handful of empirical studies directly related to test-optional institutions exist, and they are discussed below. In addition, some individual institutions have published the results of having enacted a test-optional policy, and arguments for and against such policies are examined in the following pages. However, other related topics are pertinent to the discussion of test-optional admission, as well. For instance, identifying studies that have looked at the predictive power of the SAT and ACT are particularly important as they illustrate why or why not the exams are essential elements of the college admission process. An in-depth review is conducted below of the literature pertaining to whether or not the SAT and/or ACT are discriminatory against various groups of students. It is also important to review the role of test preparation insofar as it impacts ethnic/racial disparities in SAT and ACT scores. All of these subjects will be discussed in the following as a means of covering a wide berth of the existing literature pertaining to a discussion of test-optional admissions.

Predictive Power of the SAT and ACT

The ultimate justification touted by both the College Board and ACT, Inc. organizations for why their tests are so essential to the college admissions industry is that these test scores are meant to predict a student's potential for academic success in college. They are supposed to enable admissions personnel to evaluate multiple applicants on an even playing field since the American secondary education system differs so greatly across the country. No common

curriculum is taught in all high schools nor is a uniform rubric or grading scale used to evaluate student success at the secondary level (Common Core FAQ, 2014; Common Core State Standards Initiative, 2020).[2] The SAT and ACT are each intended to provide a metric by which colleges can measure students in an equitable way while also determining how likely each applicant is to be successful at their institutions. The question, therefore, is how well do these standardized tests predict student success in college? Apparently, the verdict is split on this issue as some empirical studies have shown the tests to be valid for this purpose, whereas others have invalidated this hypothesis.

Each time the SAT has been revised by the College Board, a number of scholars have attempted to evaluate its effectiveness and validity. One such study compared the results from the 1994 cohort to the 1995 cohort of SAT takers in terms of their first-year college grade point averages (FYGPA) (Bridgeman et al., 2000). The test-takers used as the sample for this study were important because, as mentioned above, the SAT was redesigned during this time, so the 1994 contingent was the last to take the old version of the test while the 1995 group was the first to take the new one. The authors concluded that the test did an adequate job of predicting FYGPA and that the restructuring of the exam had not negatively impacted the test's ability to do so. A similar study found that even after removing what the authors termed "the *g* factor" (the underlying variance within a set of standardized tests) from the equation, both the ACT and SAT still predicted college GPA reasonably well (Coyle & Pillow, 2008).

[2] In June 2010, the Common Core Standards were published by a group of educational and political leaders from 48 states, the District of Columbia, and several U.S. territories. The Common Core Standards are the first nationwide attempt to implement a common set of English/language arts/writing and math standards for K–12 education, and they were developed with the intention of providing a uniform set of educational standards to prepare students for college. As of October 2020, 41 states, 4 U.S. territories, and the District of Columbia had ratified the Common Core (Common Core State Standards Initiative, 2020).

Much empirical research on the validity of the SAT and/or ACT to predict college success ultimately concludes that while in some contexts the exams themselves can be reasonably insightful, they are much more powerful when coupled with other factors especially high school grade point average (HSGPA) (Geiser & Santelices, 2007; Mattern et. al., 2011; Syverson, 2007; Zwick, 2007a). Using a sample of almost 80,000 freshmen enrolled in the University of California system between 1996 and 1999, Saul Geiser and Maria Veronica Santelices (2007) concluded that HSGPA was a better predictor of cumulative 4-year college GPA than SAT scores, and furthermore that the best results were found when using HSGPA and standardized test scores together. The same combination of these two variables was most effective when predicting 4-year graduation rates, as well (Geiser & Santelices, 2007).

After the SAT was redesigned once again in 2005, the College Board published research discussing the validity of the SAT particularly in light of the recent updates made to the exam's structure and scoring methods. One study examined the results of the three sections of the test (Critical Reading, Math, and Writing) in relation to how well each correlated to FYGPA. It concluded that of the three, the Writing section performed the best, and that HSGPA and SAT scores are most effective in predicting FYGPA when used in conjunction with one another (Kobrin et al., 2008). In short, even the test's parent company has recognized that the SAT is not wholly predictive of first-year college success on its own but is most advantageously used in combination with other aspects of the freshman college application package (Kobrin et al., 2007; Kobrin et al., 2008).

Comparing three particular studies, one at a large public institution, one at a small liberal arts institution, and the third at a private research institution, proves to be incredibly interesting as each came to similar conclusions. An analysis performed on the data of the University of

Georgia's freshman class in 2006 revealed that the SAT contributed little to estimating future college success, whereas high school GPA and to a small extent AP credits were better determinants (Cornwell et al., 2012, p. 135).[3] Similarly, the analysis of 17 years of college student data at a single liberal arts institution concluded that the SAT assisted in predicting college GPA by a nominal amount of three percentage points and only when coupled with other personal and academic data points (Rask & Tiefenthaler, 2012, p. 125). Lastly, a study of the freshmen at Johns Hopkins University from 2006–2008 revealed that HSGPA was again the best estimator of a student's GPA in college and that the SAT contributed marginally to the results of such estimations (Wonnell et al., 2012, p. 147).

Regarding how viable the ACT is in terms of predicting college success particularly in comparison to the SAT, Allen and Sconing (2005) argued that the ACT is well suited to determine whether or not a student is ready for college level work because the content of the ACT exam covers specific academic areas taught in high school and also correlates with particular subject areas often studied in the first few semesters of college. Furthermore, Allen and Sconing (2005) noted that ACT scores are in fact better predictors of academic performance in college than high school GPA because the ACT exam seen by all test-takers is exactly the same, whereas GPAs are calculated on different scales and are highly variable in high schools across the country.

Whether used independently or in conjunction with other factors such as high school GPA, some researchers have proven through their work that the SAT and ACT exams are not

[3] Interestingly, but somewhat unsurprisingly, high school location was also found to affect the results of the study significantly especially for Black students, which the authors noted as "…potentially troubling, because blacks [*sic*] in Georgia are overrepresented in districts with lower average school quality" (Cornwell et al., 2012, p. 135).

predictive of college success at all (Perez, 2000; Waugh et al., 1994). In some cases, it was determined that student achievement on the SAT or ACT was negatively related to satisfactory academic progress in college, indicating that the students with the lowest test scores fared the best in college while the opposite was true for those with the highest test scores (Micceri, 2010). In addition, demographic and familial characteristics impact FYGPA so much that after controlling for these factors the SAT only predicts 3% of the variance in student college success (Buchmann et al., 2010).

Criticism of the Tests as Discriminatory

Why is it that the SAT and ACT exams, which have been such an ingrained part of American college admission practices for decades, are suddenly under such attack in the past several years? Many researchers have argued that standardized college entrance exams unfairly disadvantage underrepresented minority and low-SES student populations (Espenshade & Chung, 2012, p. 177, pp. 181–182; Mattern et. al., 2011; Murray, 2012, p. 72; Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Rosner, 2012, p. 104; Syverson, 2007; Zwick, 2007a). Students from affluent backgrounds, or at least who attend high schools in prosperous areas, are more likely to be exposed to test preparation programs and materials. Their families can often pay to send them to test preparation classes, purchase preparatory software and books, or even pay an independent test preparation counselor to mentor them on a one-on-one basis; all these forms of coaching have been shown to increase a student's test scores on the SAT and ACT (Robinson & Monks, 2005; Rooney & Schaffer, 1998; Zwick, 2007a). Furthermore, most colleges and universities consider just the highest of a student's scores on the ACT or SAT for admission purposes so applicants often benefit from taking either exam multiple times as opposed to taking them just once. Therefore, students whose families can afford to pay for them

to sit for the tests several times have a distinct advantage over poorer students who may struggle to pay for even one opportunity.

In an attempt to answer the question of whether or not the exams are biased against certain groups of students, studies have looked at the performance of females versus males, various racial and ethnic groups, and socioeconomic status brackets, as well as students for whom English is not their first language (Kobrin et al., 2007; Micceri, 2010; Perez, 2002), sometimes looking even further at subgroups within groups such as the gender breakdown within a particular race or ethnicity. Across the board, the existing body of research shows that women score lower on average on both the SAT and ACT than men (Buchmann et al., 2010; Perez, 2002; Rosner, 2012, p. 106). The College Board suggested that one reason for this may be that since the 1970s the number of female students and particularly minority female students taking the SAT each year has been on the rise, which has been diluting the average scores of their subgroup (Kobrin et al., 2007). They also pointed out that men tend to achieve higher average scores than women on other common standardized tests used in the United States including the MCAT, GRE, GMAT, and LSAT (Kobrin et al., 2007), implying that this phenomenon appears to be a pervasive issue across the standardized testing industry. The results of a similar study conducted by Micceri (2010) point to the discriminatory nature of the exams and suggest not using standardized test scores for admission purposes for any female applicants at all.

Between 1987 and 2006, the percentage of SAT takers who were bilingual rose from 8 to 13 percent in the U.S. (Kobrin et al., 2007). Students whose native language is not English naturally face a disadvantage when sitting for an exam given in the English language. It is therefore unsurprising that students in this category tend to score lower on these two tests than their native English-speaking counterparts (Perez, 2002). Not only must they first interpret the

questions on the test in terms of language, but they must then also decipher what is actually being asked and provide an answer in their non-dominant tongue (Kobrin et al., 2007). Since the exam then essentially becomes a test of the student's grasp of the English language rather than of the concepts and knowledge they are designed to assess, it is expected that some wrong answers provided by these students will be the result of misunderstanding the questions themselves and not because the student did not know how to correctly answer them.

Probably the most well-known and most commonly examined aspect of the SAT/ACT exam discrimination issue against minorities is that of the disparity of average scores among various racial and ethnic groups. When separated out by racial/ethnic group, performance on the SAT follows a clearly hierarchical pattern of White students scoring the highest followed by Asian Americans, American Indian or Alaskan natives, Hispanic/Latino students, and African Americans scoring the lowest (Kobrin et al., 2007). White students tend to score higher on the SAT and ACT than students of all other races and ethnicities (Buchmann et al., 2010; Card & Rothstein, 2007; Kobrin et al., 2007; Perez, 2002; Walpole et al., 2005), and White students score 2/3 of a standard deviation higher than Hispanic/Latino students on the SAT (Rosner, 2012, p. 110). The largest differential in average SAT scores occurs between Black and White students followed by Hispanic/Latino and White students (Kobrin et al., 2007; Walpole et al., 2005). African American students' SAT scores have been found to fall between 0.82 and 1.18 standard deviations below White students' scores (Kobrin et al., 2007). The average SAT score for students at all-White high schools runs about 250 points higher than the average at all-Black high schools in the United States (Card & Rothstein, 2007).

This "Black–White gap" in both test scores and overall academic achievement has been a long-standing topic of discussion since the 1960s (Lee, 2002). An interesting study by Lee

explored this phenomenon, as well as that of the relationship between the performance of Hispanic/Latino and White students; this latter issue certainly exists but has not been studied at nearly the same level as the Black–White gap (Lee, 2002). Lee found that between the 1960s and the 1980s the disparity between academic performance of White and Black students as well as White and Hispanic/Latino students both shrunk steadily, but in the past 25 years each gap has begun to grow once more. However, the researcher did not provide a viable explanation for this trend reversal and called for other researchers to extend his line of inquiry into their own future studies.

African American and Hispanic/Latino students face additional obstacles to attending college than others because they are more likely to attend under-resourced high schools with teachers that are not as qualified, are disproportionately from urban areas, are more likely to come from low-socioeconomic backgrounds, and often deal with much more financial difficulty in paying for higher education (Walpole et al., 2005). These combined factors lead to a lower rate of participation in postsecondary education than those of other ethnic or racial groups (Walpole et al., 2005). Their historically lower performance on the two most common college entrance exams then becomes just one of the many barriers to college access that these students must overcome.

Further, it is believed that minority high school students often succumb to what is known as “stereotype threat” (Kobrin et al., 2007; Walpole et al., 2005). Stereotype threat, a concept developed in the mid-1990s, occurs when a targeted group becomes aware of a particular stereotypical belief about them and allows it to influence their performance in the activity at hand; as a result, individuals in the group end up living up to that exact expected behavior (Kobrin et al., 2007; Steele, 1997; Steele, 1998; Steele & Aronson, 1995; Walpole et al., 2005).

It is not enough that these Hispanic/Latino and African American students face additional obstacles inconceivable to their majority counterparts, but they also feel even more pressure to perform well brought on by the belief that society already thinks they will do poorly. When they do indeed end up scoring below par on the SAT or ACT as a result of letting knowledge of the stereotype impact them, it serves as reinforcement for the idea that they were not good enough in the first place. All aspects of this dilemma serve to perpetuate the cycle of minority underrepresentation in higher education.

Walpole et al. (2005) examined the way Hispanic/Latino and African American students from southern California perceived the role of ACT and SAT scores in the college admission process as well as their experiences with preparing for and taking the exams. The insights gained from interviewing these students are remarkably informative. The study found that when they were aware that the tests played a crucial role in their college aspirations, the students dealt with a critical level of anxiety about the exams as well as feelings of pressure to do well. Many students in this first group took one test or the other more than once in the hopes that their scores would improve with each sitting. On the other hand, a good number of participants seemed unaware of the important role the ACT and SAT play in the college admission process. The lack of information all these students possessed about how to prepare for the SAT/ACT, how important their scores may be to their college applications, and how their academic opportunities up until that point had prepared (or not prepared) them for that moment was obvious.

Even more concerning was the fact that most students interviewed seemed completely unaware that any college options existed that might not emphasize or require SAT or ACT scores at all as part of the freshman application package, which may have alleviated some of the pressure they felt to over-perform on the exams. The findings support the idea that stereotype

threat played a role in this situation as many minority students tend to score lower on the SAT or ACT than their counterparts. The results are even worse for Hispanic/Latino and African American males. Not only do students in these two subgroups achieve lower average scores than their counterparts but have also shown that their SAT scores actually over-predict their potential success in college (Kobrin et al., 2007). The College Board acknowledges that these disparities exist but counters this fact with the argument that the differentials between racial and ethnic populations, though still present, have been decreasing over time. The organization posits that, should its standardized test in fact be biased against certain groups, the expected result would be for minority students to achieve greater academic success in college than predicted by their test scores, which the authors claim is not the case (Kobrin et al., 2007).

The disparity among average test scores extends beyond gender, language, and race/ethnicity to socioeconomic status (SES). Previous research has documented that lower income students tend to achieve lower scores on the SAT and ACT than higher income students (Buchmann et al., 2010; Geiser & Santelices, 2007; Kobrin et al., 2007). On average, a 100-point difference can be observed between the scores of students who hail from households with less than $30,000 of annual income and students who come from households with more than $100,000 of annual income (Kobrin et al., 2007). Not only have standardized test scores been found to be highly positively correlated with socioeconomic status, but they have also been shown *not* to be correlated with high school GPA (Geiser & Santelices, 2007). This information reinforces the argument that HSGPA is indeed a much more reliable indicator of potential college success than scores on either the ACT or SAT. Students from less affluent families and communities already face many struggles in attending college, not least of which are limited resources in elementary and high schools, difficulty in financing their educations, and the

possibility of being exposed to high-crime neighborhoods, among others. If income is so closely correlated to test scores, then using SAT/ACT scores as a means of entrance to college propagates the cycle of creating barriers to college access for low-income students.

Certainly, colleges and universities choosing test-optional admission policies are one reaction to the increasing level of debate about the fitness of both the ACT and SAT to the college admissions industry. However, some researchers have stopped short of proposing a complete abandonment of the existing college entrance exam system by suggesting possible alternatives. One such proposal suggested by Roy O. Freedle (2003) recommends that the College Board simply change the way the SAT is scored. Freedle further explained the difference between *statistical bias,* which occurs when one group performs consistently poorer on the SAT than another, and *cultural bias,* which occurs when individuals of differing backgrounds interpret the same questions on the exam in different ways. Both of these phenomena lead to disparity among test scores for minority populations and their majority counterparts (Freedle, 2003). Freedle outlined a new way of scoring the SAT exam essentially omitting the easiest questions on the test from the equation because those are the ones that tend to be answered differently the most often by students of various racial/ethnic groups. Leaving in just the hardest questions brings average SAT scores of separate ethnic and racial groups much closer together. Freedle stated that, should his proposal be implemented, the average scores of minorities would increase as would their rate of acceptance at more selective colleges. This type of innovative thinking and deep evaluation of the SAT exam, along with similar assessment of the ACT, are what is needed in order to definitively determine whether or not these exams truly have a necessary place in college admissions in the 21st century.

The Role of Test Preparation

Apart from being judged as discriminatory, both the ACT and the SAT have been additionally criticized for being highly "coachable" (Buchmann et al., 2010; Perez, 2002). Indeed, an incredibly lucrative industry has developed around the idea of teaching students how to successfully navigate these exams (Briggs, 2001; Devine-Eller, 2012; Powers, 1993; Walpole et al., 2005). Companies such as the Princeton Review, Khan Academy, and Kaplan, Inc. earn millions of dollars each year selling test preparation materials and teaching courses to prepare students for the tests; in 2009, the Princeton Review made over $110 million in profit from these efforts (Buchmann et al., 2010). Test preparation takes many forms including but not limited to: review books, practice tests, private academic tutors, information concerning test and question format, coaching on test-taking strategies, and skill-building workshops on decreasing test anxiety (Powers, 1993). In more recent years, test preparation has become widely available online in addition to in-person opportunities, which increases the rate at which individuals participate (Briggs, 2001). As of 2012, it was documented that approximately 25% of high school students take part in formal test preparation endeavors annually (Devine-Eller, 2012).

The issue many people have with the idea of preparing students to take the SAT or ACT in this way is that much of the content of the test preparation materials centers upon strategies and hints for earning the highest score possible simply by understanding the way the tests are formatted and scored rather than reviewing academic material that may actually be covered in either test's questions. The College Board and ACT, Inc. are equally as reluctant to admit that their tests are coachable, yet the companies that offer courses, books, and online materials usually guarantee higher scores—sometimes as high as a 100-point increase on the SAT or equivalent jump on the ACT—as a result of an individual participating in one of their programs

(Briggs, 2001; Powers, 1993). Claims such as these illustrate why the test-preparation industry is such a lucrative one, yet "[n]o study published in a peer-reviewed journal shows average gains approaching the fabled 100-point and 200-point jumps reported anecdotally" (Murray, 2012, p. 76).

The incentive to partake in these pre-test efforts, then, is to hopefully increase a student's scores thus enhancing the student's chances of being admitted to more selective colleges (Thomas, 2004). However, research has shown that the scores of students who take either exam more than once tend to increase regardless of whether or not they have participated in test preparation, making it difficult to justify the claim of causation between test preparation and higher scores (Briggs, 2001; Powers, 1993). Other studies have also found mixed results in terms of test preparation effects (Briggs, 2001; Powers, 1993). It is therefore very easy for commercial test preparation companies to cite only those studies that prove the point they are trying to make and omit the others, essentially leading to false advertising on their parts when claiming they can guarantee a certain change in a student's scores as a direct result of taking part in their programs (Powers, 1993).

One author astutely pointed out two flaws in the philosophy behind this entire industry. First, if test preparation does indeed help students increase their scores, but it is not equitably accessible to all, then it brings an inherent aspect of unfairness to the exams. Second, if test preparation does in fact assist students in increasing their scores due not to reviewing academic content but by teaching test taking strategies, then neither exam is actually evaluating what their parent companies say they are (Powers, 1993). If either or both of these hypotheses are true, they strongly support the argument not only against the idea of test preparation but against using standardized test scores for freshman college admissions at all.

This concept of test preparation falls under the umbrella of what some term "shadow education" (Buchmann et al., 2010; Devine-Eller, 2012). Shadow education consists of all the various academic support activities that take place outside the school building aimed at improving students' academic achievement, and this practice is engaged in frequently in various countries throughout the world (Devine-Eller, 2012). However, some believe that shadow education increases the inequalities already felt across social strata in the United States when it comes to education (Buchmann et al., 2010; Devine-Eller, 2012). The connection between test preparation and the discrimination against students from low-socioeconomic backgrounds is not difficult to perceive. Test preparation is supposed to increase student scores on the SAT and ACT exams but can be extremely expensive to purchase. Students from more affluent backgrounds possess more resources with which to participate in these activities and do so at a higher rate than their low-income counterparts (Buchmann et al., 2010; Murray, 2012, p. 75) thus resulting in higher average scores for students with higher family incomes and perpetuating the bias against less affluent students in this area.

In their study on this subject, Buchmann et al. (2010) found that not only is the range of test preparation options available quite diverse in terms of format and cost, but it is directly correlated with the income level of the students who choose to participate in them; low-income students are more likely to use less expensive test preparation while high income students are more likely to choose more costly options. In addition, the types of students who engage in various levels of test preparation vary by other factors; females and those from the northeastern region of the U.S. are more likely to partake in test preparation than males and students from other parts of the country (Buchmann et al., 2010). The findings supported the previous research which found that African American and Hispanic/Latino students had much less access to test

preparation opportunities, and those that were available to them were the least expensive types often of short-term duration, certainly not of the same caliber as those to which White and high-SES students typically have access.

The research of Derek Briggs (2001) and Audrey Devine-Eller (2012) corroborated the findings of the previous study in showing that high-income students disproportionately use test preparation activities. Asian American students, those who spend more hours outside of school studying, those with educational aspirations of attending more selective schools, and those whose parents and families are more aware of the importance of the SAT/ACT in regards to college admission, are all more likely to engage in test preparation than other groups (Briggs, 2001). In addition, Devine-Eller (2012) supported that the Black–White test score gap exists between students of these racial groups even for those who come from the same socioeconomic background.

Both the ACT and SAT testing organizations offer fee waivers to allow low-income students the opportunity to take their tests for free (ACT, Inc., 2020a; College Board, 2020a) as a means of combating the barrier that the price of sitting for the tests may present to low-SES students. The ACT waiver covers a maximum of four exams (ACT, Inc., 2020a) whereas the SAT waiver covers two exams (College Board, 2020a). Indeed, students who can afford to pay for the exams tend to take them multiple times. Half of all SAT takers complete the test at least two times, and 15% take the SAT three or more times (Buchmann et al., 2010). Students who sit for the exams multiple times or who take both the SAT and ACT do so under the assumption that these behaviors will result in higher scores (Thomas, 2004). In fact, studies have shown that compared to high- and low-achieving students, those who are considered borderline (e.g., those

with high HSGPAs and low test scores or vice versa) tend to take either the SAT, ACT, or both tests multiple times with just such a goal in mind (Thomas, 2004).

Yet despite these facts, families readily contribute their hard-earned money to these entities year after year in exchange for the belief that they are giving their students every additional edge possible over their peers when it comes to the competitive enterprise of admission to a selective college. Over the years, the media-fueled frenzy over the supposed ruthlessness of the college admission process has exacerbated these perceptions. Even though average SAT scores have increased in recent decades, the admission rates particularly at selective schools have decreased (Walpole et al., 2005). This reality coupled with highly publicized stories of high-achieving students being denied entry to the colleges of their choice leads to higher than ever levels of student anxiety over ACT and/or SAT score performance (Walpole et al., 2005).

Whether or not it is true that companies offering test preparation can actually keep their promises of higher scores on the SAT or ACT, the practice of participating in test preparation is pervasive among American high school students. The cost of test preparation materials ranges from a few dozen dollars for purchasing books with which to coach oneself to several thousand dollars for in-person classes or private tutors (Briggs, 2001). Considering that the College Board charges $52 without and $68 with the optional essay section for taking the current version of the SAT (College Board, 2020b), and the ACT costs $55 without and $70 with the optional writing section (ACT, Inc., 2020c, students and their families may very well pay several times more to prepare to take the exams than to actually sit for them.

It should be noted that in an effort to even the playing field in regards to test preparation, along with the changes to the SAT test itself released in 2016, the College Board has also

partnered with Khan Academy to provide free online test preparation materials for the first time (Khan Academy, 2016). The future will reveal whether or not this initiative does indeed assist low-SES and underrepresented minotiry students in achieving higher scores on the SAT as well as how much it erodes the profitability of costly test preparation provided by the various companies mentioned above.

Justification for Test-Optional Policies

It has been well established in the extant literature that the SAT and ACT exams face widespread criticism on a variety of factors. Some say they are discriminatory and create unnecessary barriers to college access while others say they are inconclusive when determining college preparedness and are therefore pointless to the college admission industry. Whatever the argument, the concept of college entrance examinations has come under scrutiny in recent years, and many colleges and universities in the United States are beginning to respond. More than 1300 institutions of higher education in the U.S. have adopted test-optional admission policies where some if not all of their freshmen applicants are not required to submit SAT/ACT scores as part of their application portfolios (National Center for Fair and Open Testing [FairTest], 2020).

The reasons espoused by various scholars as to why colleges might choose a test-optional admission policy run the gamut of possibilities from moral to strategic. Schools may look at a test-optional policy as a vehicle for achieving other goals. For instance, a college may aim to increase its total applications (Epstein, 2009; Robinson & Monks, 2005). Students who previously felt intimidated by a particular college's reported average test scores might now choose to apply for admission knowing test scores are not required (Espenshade & Chung, 2012, p. 185; Rooney & Schaeffer, 1998). In addition, by raising its total applications yet admitting the same number of students as in prior years, an institution might decrease its acceptance rate

thereby increasing its perception of selectivity (Epstein, 2009; Robinson & Monks, 2005). Since it is a metric often used to assess quality and reputation, it is easy to see how decreasing an institution's admission rate might be viewed as beneficial.

As previously discussed, it is well documented that minority students, females, and those from low-socioeconomic backgrounds score, on average, lower on these tests than their White, male, and high/mid-SES counterparts (Buchmann et al., 2010; Dixon-Roman et al., 2013; Nankervis, 2011; Perez, 2002; Syverson, 2007; Zwick, 2007a). Therefore, by removing the barrier of requiring SAT scores in the admission process, colleges with test-optional policies are opening their doors to a more diverse applicant pool (Epstein, 2009; Espenshade & Chung, 2012, p.190; Rooney & Schaeffer, 1998; Syverson, 2007; Syverson et al., 2018). For colleges and universities for whom increasing diversity rates on campus has become a priority, implementing a test-optional admission policy might be a viable mechanism for achieving this goal as it was, for instance, at Wake Forest University beginning in 2008 (Allman, 2012, p.173, p. 175).

Continuing in the line of strategic reasons, existing research identifies purposeful score inflation as another root cause for colleges to implement a test-optional admission policy (Epstein, 2009; Robinson & Monks, 2005; Syverson, 2007). Unsurprisingly, in test-optional situations students who do well on the exams are likely to include their scores with their application packages. By not requiring all enrolling students to report test scores it can be assumed that students with the lowest scores will disproportionately choose to leave this piece of data out of their applications. However, colleges with test-optional admissions often still report their average test scores publicly, even though their calculations essentially only use partial data, and in this way these institutions could be seeing increases in their numbers. This becomes particularly critical when independent organizations such as *U.S. News and World Report*

publish rankings of colleges and universities and include average or middle 50% test scores in their calculation of rank. Increasing one's reported average test scores could assist in driving up a college's ranking on such a list (Syverson, 2007). This is not only significant to institutions of higher education in terms of concrete rankings but also in less tangible means such as reputation and public image as a selective school (Syverson, 2007).

Beyond the reasons for going test-optional for reputational or calculable purposes, there are arguments related to the true purpose of a college admission process. Among others, both Rebecca Zwick (2007a) and Steven Syverson (2007) documented that SAT scores can predict academic success in college to a certain extent but *only* in combination with other key factors such as high school performance and grade point average. Nevertheless, Zwick (2007b) indicated that many in the admissions field feel the SAT is too narrow of a metric to be a true indicator of potential college success. Marilyn Gilroy (2007) cited the former dean of admissions at Drew University in New Jersey who stated that internal research proved high school GPA to be the most important variable in predicting college performance for that school's applicant pool. Thus, some colleges are removing test scores from the admission equation simply because they feel the other pieces of a student's application provide ample evidence for or against admitting the student (Gilroy, 2007; Robinson & Monks, 2005).

Yet the several proposals discussed thus far are not sufficient to cover all the reasons why an institution of higher education might adopt a test-optional admission program. Some colleges feel obligated to offer an admission process that falls clearly under the school's mission, as is the case for Providence College (Shanley, 2007). Providence College, located in Rhode Island, has had a tradition of being committed to providing higher education opportunities to minorities and students of low-SES backgrounds since its founding in 1917. As such, opening its doors to

diverse, economically disadvantaged and underrepresented groups is embedded in Providence's admission process for first-year students. For the many reasons previously discussed regarding the unequal results on the SAT for students who fall into just such categories, Providence moved to a test-optional program in 2006 (Shanley, 2007). The administration of Providence College views its mission to be more important than reputation or ranking and has remained part of the test-optional contingent since launching its pilot program at that time. In a similar vein, St. John's College in Maryland (Syverson, 2007) and Muhlenberg College in Pennsylvania (Rooney & Schaeffer, 1998) both have gone test-optional for reasons of renewing a holistic admission process as well as returning the admission focus to selecting the student that is the right fit for a particular college and vice versa.

The Case Against Test-Optional Policies

The research also discusses the contrary argument, namely reasons not to go test-optional. The primary argument for such a policy is that by removing standardized test scores from the admission equation, the academic quality of a college's incoming first-year class will decrease (Rooney & Schaeffer, 1998; Syverson, 2007). Critics of test-optional policies also argue that a school's reputation and image as a selective school will suffer (Rooney & Schaeffer, 1998; Syverson, 2007). Furthermore, Jonathan P. Epstein (2009) contended that it is less fair to prospective students of a given college or university to consider the test scores of only some of their applicants than it is to consider them for either all or none. Allowing the applicant rather than the institution to determine if SAT scores will be part of the review process is illogical and unjust since there is no consistency to this practice.

From a more pragmatic point of view, some admission offices believe that SAT scores provide a worthwhile metric on which to evaluate applicants and that these scores add real value

to the admission process. They fear that removing test scores from admissions will prevent them from properly admitting a successful incoming class (Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Syverson, 2007). Rooney & Schaeffer further argued that standardized college entrance exams, as imperfect as they may be, at least offer a "common yardstick" (p. 10) against which all applicants to a particular school may be uniformly measured.

Test-Optional Policy Outcomes

Several articles in the available literature present case studies of the experiences of individual institutions, and many articles in the mainstream media cited particular statistical results of adopting said policies. However, few large-scale empirical analyses have been performed on the characteristics of institutions that tend to choose test-optional admissions or on the outcomes of such policies in terms of enrollment, retention, or graduation; furthermore, there is little research on the change in diversity rates (Belasco et al., 2015).

Although the results vary in size and scope, most colleges that are test-optional report similar findings in several areas. Most noticeable of all is the increase in total application volume; this phenomenon was observed by a host of institutions including Bates College (Epstein, 2009; Hiss, 2011), Providence College (Epstein, 2009), and Mount Holyoke College (Robinson & Monks, 2005). Multicultural, economically disadvantaged, and first-generation students are highly represented among the students making up the additional applicants over prior years (Epstein, 2009; Robinson & Monks, 2005; Syverson, 2007). All these colleges reported no significant difference in the academic quality, college performance, or graduation rates of the incoming freshman students between those who were admitted with SAT scores and those admitted without them (Epstein, 2009; Robinson & Monks, 2005; Syverson, 2007). These

outcomes provide strong credence to the arguments for adopting test-optional admission programs.

Charles Rooney and Bob Schaeffer (1998) presented the cases of Franklin & Marshall College, Bates College, and Muhlenberg College, as well as the institutions in the California State University system and the Texas public university system. Using these five examples, Rooney & Schaeffer concluded that SAT and ACT scores provide little additional value to an application and that high school academics are the best indicator of gauging potential college success. The findings also suggest that test-optional colleges and universities enjoy increased diversity among their applicants. It should be noted, however, that this particular study was produced with the support of the National Center for Fair and Open Testing, and therefore it cannot be stated definitively that the study was conducted independent of any influence from said organization (Rooney & Schaeffer, 1998).

Similar to the results shown by the five case studies above, Shanley (2007) reported Providence College's positive test-optional experience. Fr. Shanley heavily stressed Providence's dedication to providing higher education to traditionally underserved populations and cited this as Providence's primary reason for following the test-optional trend. The school's internal research showed that high school grade point average had historically been a better predictor of college success than test scores, and officials at the institution feared that they were missing out on the chance to recruit excellent students who were deterred by Providence's published average SAT scores, leading to the decision to de-emphasize test scores in admissions. In its first year after making this decision, Providence College observed a 20.8% increase in applications from racial/ethnic minorities and a 21% increase in first-generation applicants. These changes carried over to substantial increases in the percentage of minority and first-

generation students who enrolled at Providence College in the fall 2007 semester as well as those in the lowest socioeconomic brackets as evidenced by their eligibility to receive a Pell Grant (Shanley, 2007).

Belasco et al. (2015) conducted the first large-scale empirical study examining questions surrounding the effects of test-optional policies upon a group of institutions. Belasco and his colleagues questioned whether test-optional institutions truly enjoy greater diversity, defined by low-socioeconomic status and ethnic/racial minority status, than their counterparts that require SAT/ACT exams. They concluded that the diversity seen on test-optional college campuses was not in fact significantly different from test-requiring institutions nor had it grown proportionally over the 20 years studied. Rather, the assumed selectivity of test-optional colleges and universities benefited from having both higher average SAT scores and greater application numbers (Belasco et al., 2015).

Syverson et al. (2018) used data from 28 institutions in a peer-matching format to determine the outcomes of test-optional policies. By comparing individual test-optional schools to specific test-requiring peer institutions of similar size, control, and underrepresented minority and Pell Grant recipients enrolled, the authors determined that test-optional programs increased applications and enrollment of low-income and racial/ethnic minority students (Syverson et al., 2018).

Although many have written on the advantages and disadvantages of test-optional admission policies, few discuss possible alternative methods of pursuing unbiased college admissions. However, at least two alternatives to including SAT/ACT scores in the process have been suggested in the literature: percent plans and lotteries (Zwick, 2007b). Each proposed method is evaluated in terms of its ability to effectively create a fair and equal-access procedure

for handling college admissions. Although each has its merits, Zwick (2007b) concluded that neither of these options are perfect substitutes for standardized tests and that none of these approaches achieve the true goals of providing equal opportunity to higher education for all students and showing true diversity in all its forms on a college campus.

Conclusion

Some authors suggest that a test-optional policy is not necessarily feasible to implement or easy to achieve in all types of institutions. When SAT scores are removed from the equation, other elements of the application must be examined in more depth. Smaller schools with a manageable applicant pool relative to the size of their admission staff often possess the time and resources to delve deeply into each and every application, sometimes multiple times. This is the type of admission environment where a test-optional policy thrives (Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Syverson, 2007; Zwick, 2007b). Large flagship public universities receiving thousands of applications per year traditionally require some type of objective metric by which to pare down their applicant pools on the first pass, which is how SAT and ACT scores come into play, for example. Were large institutions to eliminate standardized test scores from their application review procedures, they might require the addition of so many more admissions staff members as to negate the advantage of making such a policy change (Schaeffer, 2012, p. 165). Should this in fact be true, the implication is that this could greatly impact the discussion currently surrounding the question of test-optional admissions. Studies such as this seek to illustrate whether or not there are similarities in outcomes among test-optional institutions regardless of size or type, but perhaps even if there is evidence of positive results it still will not be feasible to implement test-optional admissions at every institution that might desire to do so.

Chapter III

Methods

Overview of the Study

This study aimed to present empirical research on the outcomes of test-optional admission policies at 4-year colleges and universities in the United States. The topic of test-optional admissions is closely related to the idea of increasing access to higher education, as many in the education field believe standardized tests such as the SAT and ACT are discriminatory toward minority and other underrepresented groups and therefore act as a barrier to college access for these types of students (Espenshade & Chung, 2012, pp. 177, 181–182; Murray, 2012, p. 72; Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Rosner, 2012, p. 104; Syverson, 2007; Zwick, 2007a). In addition to contributing to the policy discussion of access, this study sought to provide empirical evidence on how particular outcomes at colleges and universities differ at test-optional institutions versus those that require SAT or ACT exam scores.

Research Questions

This study asked three key research questions:

1. Are the diversity rates of test-optional institutions significantly different from those of test-requiring institutions, as measured by the proportion of racial and ethnic minority students and Pell Grant recipients enrolled?

2. Do institutions with test-optional policies gain a competitive advantage in reputation and selectivity as measured by a decrease in acceptance rates and/or an increase in national rankings in the *U.S. News and World Report*?

3. What impact, if any, does the adoption of a test-optional admission policy have on student success as measured by first-year retention and 6-year graduation rates?

Conceptual Framework

In terms of a conceptual framework, the premise of this study is grounded in two distinct and yet interconnected theories, each of which correspond to the particular research questions at hand. First, to provide some background for the first and third questions regarding the impact of test-optional admissions policies, it is necessary to look into the intended and unintended consequences of adopting (or not adopting) such a policy. Robert K. Merton's (1957) theory of latent and manifest functions provides a theoretical basis for further understanding this phenomenon. To address the second research question regarding competition in higher education, Shelby Hunt and Robert Morgan's (1995) theory of comparative advantage of competition provides a framework in which one can understand institutional behavior in this regard at a deeper level.

First, rooted in the field of sociology, Merton's (1957) manifest functions refer to "...those objective consequences contributing to the adjustment or adaptation of the system which are intended and recognized by participants in the system" whereas latent functions are "neither intended nor recognized" consequences (p. 51).

Merton's (1957) concept of latent and manifest functions is interestingly juxtaposed with the idea of test-optional admissions. It can be argued that in the field of college admissions, adopting a test-optional policy has the clear manifest, or intended, consequence of increasing an institution's access to students of more diverse economic and racial/ethnic backgrounds. It could also have the manifest function of allowing institutions to distance themselves from the growing stigma of the SAT and ACT as they become increasingly criticized for their alleged discrimination against minority students. Various other objectives of becoming a test-optional school would all fall under the category of manifest functions, too, as long as the institution

listed the outcome as a stated goal from the outset. Some of these manifest functions might be outwardly stated while others are kept private, but nevertheless if they were intended outcomes before the policy change, then they fall into the manifest function category.

As anyone who has embarked upon a new project or policy is aware, even with the best intended plans, there might be unintended consequences that arise in addition to the expected ones, and these are what Merton (1957) called latent functions. In terms of test-optional admissions, an example of latent functions might be increasing the total number of applications overall when the intention was merely to increase the number of applications from diverse students. This could potentially lead to an additional latent function of increasing the perception of the institution's selectivity. If overall application numbers increase but the college admits the same number of students as prior years, its acceptance rate drops, which is often viewed as a measure of selectivity and institutional prestige.

Next, Hunt and Morgan's (1995) theory of the comparative advantage of competition was originally conceived as a model of competition within economics in response to the neoclassical theory of perfect competition. The authors made the case that neoclassical theory does not satisfactorily account for certain phenomena within the business world such as the diversity of organizations, among other flaws. In contrast to the neoclassical theory, the comparative advantage theory of competition acknowledges some truths about market-based economics that are unexplained by other theories, including the fact that consumers have individualized preferences that are continuously in flux and that these consumers are entering the market economy with imperfect or partial information about what products and services are available to them (Hunt & Morgan, 1995). Additionally, the primary job of managers within an organization is to identify, analyze, implement, and assess strategies that will lead their organization to a place

of advantage above its competitors in the marketplace. The core objective of the firm, then, is to reach "superior financial performance" (Hunt & Morgan, 1995, p. 6).

The crux of the comparative advantage theory of competition is the premise that organizations are continuously gaining a competitive edge over their peers, often achieved by obtaining a resource that their competitors do not have, allowing them to produce a higher quality good or service at a lower price.

> Once a firm's comparative advantage in resources enables it to achieve superior performance through a position of competitive advantage in some market segment or segments, competitors attempt to neutralize and/or leapfrog the advantaged firm through acquisition, imitation, substitution, or major innovation. (Hunt & Morgan, 1995, p. 8)

It seems that, although the higher education industry has traditionally been treated as separate from the business world, certain concepts regarding consumer and organizational behavior in economics can be applied to higher education. Colleges have marketing departments and promote their own positive messaging in the marketplace similar to businesses. By viewing students as "consumers," colleges and universities with leadership that understand the admissions industry are keenly aware of who their peer institutions are: that is, institutions with whom they are most often competing for students. Frequently these peer institutions are similar in size and scope, academic offerings, selectivity, and sometimes geographic location. College admissions professionals are continuously looking for ways to market their own institution as unique and to highlight the ways in which their college stands out among its peers in order to entice students to first become interested in learning more, then to apply for admission, and eventually to enroll.

Taking the comparative advantage theory a step further, obtaining a resource that competitors do not possess provides an institution with a means of gaining an edge in the marketplace. I would argue that a test-optional policy could be seen as the specific resource in

question. As aforementioned in the literature review section, institutions of higher education are often quick to state that their intention when going test-optional is to enhance diversity on campus, increase access to college for underrepresented populations, and/or return to a holistic view of admissions. However, some colleges might also be looking for a way to meet more self-serving goals, which they are much less likely to admit publicly, such as raising their reported average SAT and/or ACT scores or lowering their admission rates, commonly viewed in the industry as a measure of selectivity. By becoming a test-optional institution, a college may hope that the results will advantage them in ways both explicitly and implicitly stated as the primary objectives. Additionally, it could also be argued that becoming a test-optional institution might be viewed as the innovation that colleges are bound to develop in response to peer institutions gaining a competitive edge of another kind.

Together, Merton's (1957) theory of latent and manifest functions as well as Hunt and Morgan's (1995) theory of the comparative advantage of competition provide a conceptual framework that guides the empirical analysis of this study. By bringing together theories from sociological and economic disciplines, this study frames the competing institutional interests of diversity and competitive edge in the context of test-optional policy, an emerging structural change in college admissions.

Data Sources

Integrated Postsecondary Education Data System

Even though the United States does not have a national higher education system, the federal government still plays an important role in administering a fairly robust financial aid program and providing funds to college students through various grant and loan programs. As a

stipulation of participating in said programs, more recent amendments to the Higher Education Act of 1965 specify that all colleges and universities that participate in federal financial aid programs must complete a detailed set of surveys each year of various institutional characteristics and statistics (National Center for Education Statistics [NCES], 2018). The National Center for Education Statistics (NCES) collects and maintains these data in its Integrated Postsecondary Education Data System (IPEDS) database and makes the information publicly available online. According to the NCES, the IPEDS database is populated with data from more than 7,500 institutions each year (National Center for Education Statistics [NCES], 2018). Because the focus of this study was on investigating the institution-level impacts of test-optional admissions policies, IPEDS was utilized for constructing a dataset.

U.S. News and World Report

The majority of data in this study were obtained from IPEDS with the exception of the data on institutional rankings. For institutional rank data, historical copies of the ranking lists published annually by *U.S. News and World Report* (USNWR) were obtained and cataloged manually into the dataset. USNWR annually publishes several lists of college rankings drawing from the Carnegie Classification's Basic Classification system to group institutions into four categories. In recent years these lists have been titled National Universities, National Liberal Arts Colleges, Regional Universities, and Regional Colleges[4], but in the past the latter two have sometimes been called Universities - Master's and Comprehensive Colleges - Bachelor's, respectively. USNWR calculates its rankings based upon values across 17 data points (Morse & Brooks, 2020), assigning each institution a total score and then ranking them in order on their

[4] The *U.S. News and World Report's* Regional Universities and Regional Colleges ranking lists are further broken down into four geographical subgroups: North, South, West, and Midwest.

respective lists with the institution(s) with the highest total score receiving a rank of 1 in each of the four categories.

U.S. News and World Report's website, which describes its methodology for calculating their rankings, explains that only institutions in the top three quarters of each list are published in order of rank and with their total scores. Those colleges that fall into the bottom 25% of each list are published in alphabetical order with neither a numerical rank nor a total score printed alongside them (Morse & Brooks, 2020). Because of this, as well as the fact that there are four separate ranking lists to reconcile, neither absolute rank nor total score can accurately be used in this study to capture a potential change in rank over time. Rather, one indicator that is published for most institutions that can be used to better assess change over time is the peer assessment score.

The peer assessment score, one of the metrics included in the USNWR ranking calculations, is a numerical value on a scale of 1.0 to 5.0 captured by surveying top officials at colleges and universities across the country on their perception of the academic quality and reputation of other institutions. *U.S. News* maintains that this metric captures intangible qualities that cannot otherwise be quantified in order to include them in the overall ranking calculations (Morse & Brooks, 2020). Because this value is published for more institutions than rank and due to the fact that this metric is universally reported on the same scale across all lists, this is the variable that was used as a proxy for rank in this study.

Sample Selection

Data were collected from IPEDS and USNWR for the academic years 2008-2009 to 2015-2016. Out of more than 7,500 institutions in the IPEDS database, the sample was first restricted using the IPEDS data center filters to only degree-granting, 4-year institutions in the

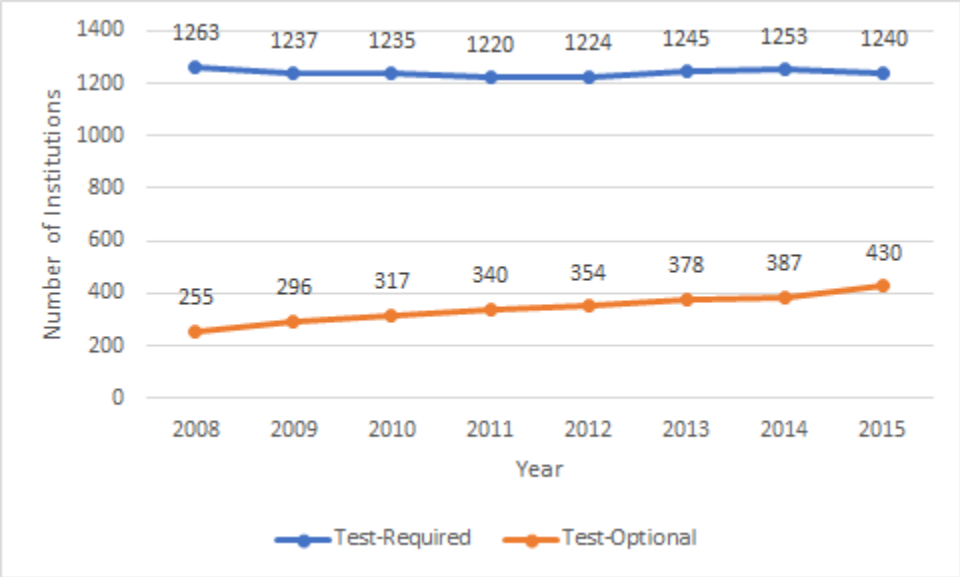
United States. After the set of institutional variables was downloaded from IPEDS and imported into SPSS (Version 26), additional restrictions were placed on the sample. Even though the filter for selecting only 4-year institutions had been placed on the IPEDS download, some 2-year and primarily associate's-degree-granting institutions remained in the sample. These and other outliers were removed from the sample by deleting all institutions from the dataset except those in the Carnegie Classification groups for baccalaureate, master's, doctoral, and special focus 4-year institutions, which removed 1,502 institutions from the dataset. The sample was then further limited to only those institutions that specified in the IPEDS data that they did not have an open admissions policy, removing an additional 889 institutions. Upon further examination of the data in the sample, I observed that 35 schools in the dataset resided in various U.S. territories rather than one of the 50 states plus the District of Columbia. However, all of these institutions were missing data for the admissions requirements variables and therefore the state of their SAT/ACT policy for freshman applicants could not be determined. They were thus eliminated from the dataset. All of these restrictions in selecting cases to remain in the dataset resulted in a final sample of 1,681 colleges and universities.

The logic behind the decision to exclude institutions with open admission policies from the sample stems from the definition of an open admission policy. Institutions with open admission policies admit all applicants who meet basic minimum requirements; for example, community colleges often have no admission requirements besides residency in a given county or other locality. Because of this policy, standardized test scores do not play a part in the admissions process of open admission schools. Under a very loose definition, open admission schools might be considered test-optional, but because they never required scores they therefore never made the conscious choice to stop requiring scores. Thus, including these institutions in

this study would have stacked the sample of test-optional schools and skewed the results in unintended ways. The breakdown of test-required and test-optional institutions within the sample are represented in Figure 2.

Figure 2

Admission Test Score Policies of the Sample, 2008–2015



Note. Data from IPEDS (2020).

As discussed in Chapter II, the test-optional movement has been growing in recent years. As shown in Figure 2, the cohort of test-optional institutions in the sample grew steadily over each year. In 2008, test-optional colleges constituted 255 institutions, which was just 15% of the sample, but by 2015 they constituted 26% of the sample at 430 institutions.

Selection of Years in the Study

The IPEDS survey on admissions includes a set of questions regarding admission practices and required application materials for first-year applicants. The IPEDS website houses historical copies of survey questions dating back to the early 1990s allowing researchers to observe the evolution of questions asked over the years. The IPEDS question regarding freshman

application requirements has not been altered much in the past 25 years, but in 2000 a major change did occur to the answer choices provided to institutions. Prior to 2000, the IPEDS survey question read as follows: “Which of the following data does your institution use as part of the selection process for entering freshmen? Mark all that apply.” Institutions were given the choice of selecting the SAT, the ACT, or an option for “other” (National Center for Education Statistics [NCES], 2020b). These possible answers left no room for nuance and gave institutions no ability to explain any alternate testing policies they may have had apart from the option of checking the box for “other.”

On the 2000-2001 survey, however, a revision changed the wording of the admission question to “...please select the box that best describes how your institution uses any of the following data in its selection process.” One of the items listed in the provided grid of application materials was “Admissions test scores (SAT, ACT, etc.).” At this point in time, the answer choices were expanded to 1) required, 2) recommended, 3) neither required nor recommended, and 4) don’t know (National Center for Education Statistics [NCES], 2020b). By rewording the question and adding additional response categories, NCES allowed institutions to more accurately reflect their complicated test score policies. Only institutions that require standardized test scores would answer “required” to this question, leaving those with more flexible test score policies for some if not all of their first-year applicants to choose one of the other options.

On the 2001-2002 survey, the admission considerations question was again reworded to read “Please select the option that best describes how your institution uses any of the following data in its undergraduate selection process.” The question has remained in this form through the 2019-2020 survey. However, on the 2016-2017 survey, the answer choices for this question were

slightly revised once again. At this time, the possible answer choices became 1) required, 2) considered but not required, 3) recommended, and 4) neither required nor recommended; the option for "don't know" was removed at this time (National Center for Education Statistics [NCES], 2020b).

Due to the shift in the composition of the admissions test score question, which occurred on the 2000-2001 survey along with availability of data on the admissions question from IPEDS, it was determined that the starting point for the study could not be any earlier than 2000. However, once data were collected from both data sources and analyzed for completeness, it was found that information was available for the most variables[5] for the academic years 2008-2009 to 2015-2016. Thus, the years to be encompassed within this study were ultimately determined by the availability of the data.

Variables

It is interesting to ponder whether any characteristics can be determined to be shared among the institutions with test-optional policies in place or whether any specific propensity towards being test-optional or not can be identified by examining basic demographic and descriptive variables. In an effort to answer this inquiry, the variables in Table 2 were extracted from IPEDS and analyzed using descriptive statistics. Due to the nature of descriptive statistics and the fact that this study spans 8 years, it made the most sense to analyze the sample at one point in time. Therefore, each of the variables in Table 2 were analyzed using descriptive statistics for the most recent academic year in the study, 2015-2016.

[5] The USNWR peer assessment score data for 2013 was never located and therefore could not be added to the study. However, the USNWR rank variables were not imputed in the study, as discussed in the later section on multiple imputation.

Table 2

Variables Used in Descriptive Statistics

Variable	Description	Coding scheme	Data source
Carnegie Classification	Since the 1970s, the Carnegie Commission on Higher Education has classified institutions of higher education into categories grouping them roughly by comparable institutions. This variable represents the basic classification of each institution in the sample from the 2015-2016 IPEDS survey year.	1 = Associate's 2 = Special Focus 2-Year 3 = Doctoral 4 = Master's 5 = Baccalaureate 6 = Special Focus 4-Year 7 = Tribal Colleges	IPEDS
Control of institution	The control indicates whether the institution is public, private not-for-profit, or private for-profit.	1 = Public 2 = Private Not-for-Profit 3 = Private For-Profit	IPEDS
Degree of urbanization	The degree of urbanization indicates the geographic environment of each institution by various descriptions of rural, town, suburban, and city locations.	1 = City 2 = Suburb 3 = Town 4 = Rural	IPEDS
U.S. state	The U.S. state in which each institution is located re-coded into four regions based upon designations used by the U.S. Census Bureau.	1 = Northeast 2 = South 3 = Midwest 4 = West	IPEDS
Total number of applicants	The total number of first-time full-time undergraduate applicants to the institution.	1 = <=5000 2 = 5001-10,000 3 = 10,001-15,000 4 = 15,001-20,000 5 = 20,001-25,000 6 = >25,000	IPEDS
Total number of enrolled students	The total number of full-time undergraduate students enrolled in the fall semester.	1 = <=1000 2 = 1001-2000 3 = 2001-3000 4 = 3001-4000 5 = 4001-5000 6 = >5000	IPEDS
Percent females applied	The percent of first-time full-time female students that applied to the institution.		IPEDS
Percent females enrolled	The percent of first-time full-time female students that are enrolled at the institution.		IPEDS

Because there are institutions in the study located in all 50 U.S. states plus the District of Columbia, running descriptive statistics on state alone would not provide much telling information by itself. Therefore, states were grouped into the following four regions based upon the U.S. Census Bureau's definition: Northeast, South, Midwest, and West (U.S. Census Bureau, 2015). Thus, better conclusions could be drawn about whether or not test-optional or test-required schools tend to be located in certain regions of the country.

In addition to this descriptive analysis, in order to answer the three key research questions the data obtained from IPEDS and USNWR were used to run a more advanced statistical model. Each of the variables in Table 3 was used as a dependent variable in this study.

Table 3

Dependent Variables

Variable	Description	Data source
Diversity rate: racial/ethnic minority students	Percent of total undergraduate enrolled students of racial/ethnic minority backgrounds	IPEDS
Diversity rate: low-socioeconomic background	Percent of total first-time full-time enrolled undergraduate students awarded Pell grants used as a proxy for low socioeconomic background	IPEDS
Undergraduate acceptance rate	The percent of total first-time applicants admitted to the institution	IPEDS
Rank in *U.S. News and World Report*	Peer assessment score used as a proxy for rank	USNWR
Full-time retention rate	The percent of full-time first-year undergraduate students who return for a second year	IPEDS
6-Year graduation rate	The percent of undergraduate students who complete a bachelor's degree within 150% of normal time (6 years)	IPEDS

Several notes should be made regarding the structure of the variables listed above. The diversity rate as measured by racially and ethnically diverse students was calculated by dividing

the total number of non-White students into total enrolled students.[6] Similarly, the diversity rate as measured by students of low-socioeconomic backgrounds was calculated by dividing the number of first-time full-time enrolled students receiving Pell Grants[7] into the total first-time full-time enrolled students. Additionally, acceptance rate is not a variable included in the IPEDS database itself. Rather, this variable was calculated by dividing total admitted students by total applicants, both of which were obtained from IPEDS. Finally, as previously mentioned, peer assessment score was used as a proxy measure to represent USNWR rank in this study.

In this study, the test-optional admissions policy (or lack of said policy) of the institution was used as an independent variable. The IPEDS survey does not ask colleges and universities outright whether or not they have a test-optional policy. As such, the existence of this policy must be interpreted from the institutions' answers to the admission test score question outlined above. The data for this question were re-coded into a binary variable stating whether or not an institution had a test-optional policy where 1 meant "yes" and 0 meant "no." Institutions that claimed that admissions test scores were required or they did not know the policy were re-coded with 0 indicating they were test-requiring institutions, whereas institutions that stated admissions test scores were recommended or neither required nor recommended were re-coded with 1 and were therefore considered test-optional schools. Institutions that indicated test scores were neither required nor recommended were included in the category of test-optional institutions because the definition of test-optional being used here includes any alternative policy to that of

[6] The diversity rate measured by racial/ethnic minority students enrolled includes Asian American students. It should be noted that in some studies, diversity is examined by looking at metrics specifically pertaining to disadvantaged or underrepresented minority students which often does not include Asian American students, but in this case all non-White students were included in the category of racial/ethnic minorities.

[7] Federal Pell Grants are generally awarded to financially needy undergraduate students and therefore may be used as a proxy to determine low-socioeconomic status.

requiring test scores for all first-year applicants. Schools that neither require nor recommend test scores are therefore considered test-optional by this definition.

The third set of variables involved in the analysis was a group of additional predictor variables included in the model that contribute to the overall effect of the key independent variable(s) on the dependent variable(s). In other words, the analysis determines what impact the primary independent variable(s) of interest has on the dependent variable(s) while controlling for the additional independent variables. The covariates selected for this study were chosen because of their relation to the topic of admissions and enrollment, shown in Table 4.

Table 4

Additional Independent Variables

Variable	Description	Data source
Full-time undergraduates Enrolled	Total full-time undergraduate students enrolled	IPEDS
Total applicants	Total number of full-time first-time undergraduate applicants	IPEDS
Percent female enrolled	Percent of total full-time undergraduate enrolled student population that are female	IPEDS
Percent receiving grant aid	Percent of total undergraduate students receiving federal, state, local, or institutional grant aid	IPEDS
Average SAT Critical Reading plus Math score	The average of the 25th and 75th percentiles of the SAT Critical Reading score plus the average of the 25th and 75th percentiles of the SAT Math score	IPEDS

The variable for average SAT Critical Reading plus Math score had to be calculated using raw score data provided by IPEDS. SAT scores are typically reported as a total of the Evidence-Based Reading and Writing section (formerly Critical Reading section) and the Math section on a 1600-point scale. In order to compute the SAT score variable used in this analysis,

first the 25th and 75th percentiles of the SAT Critical Reading section were averaged together and the same was done for the 25th and 75th percentiles of the SAT Math score. Then the two averages were added together resulting in a single average SAT score variable.

Data Preparation

Data Cleaning

The statistical software SPSS (Version 26) was used to prepare the data for this analysis. After the data were downloaded from IPEDS in CSV format, the files were uploaded into SPSS. Label names, value definitions, and missing value labels were input into the dataset for clarity when viewing the SPSS output.

The nature of the data included in this study is such that, although all but *U.S. News and World Report* rank values were retrieved from the same IPEDS data source, IPEDS contains data from several different surveys completed annually by participating institutions (e.g., institutional characteristics, fall enrollment, graduation rates, etc.). Some survey components collect data for the same academic year in which they are submitted, whereas others request final information from prior academic years instead (Jaquette & Parra, 2014). It is therefore necessary to be cognizant of how data are combined into a single dataset from multiple surveys in order to make sure that data points from a given academic year are matched up correctly across the dataset. Variables in the dataset were accordingly labeled with the corresponding year so that data across IPEDS surveys could correctly be identified as pertaining to the same academic years.

The statistical test employed for the data analysis was a fixed-effects panel regression model, which will be discussed in more detail in the later section. In order to perform a fixed-effects panel regression, that is an analysis of data spanning several points in time, the dataset

must be converted from a traditional (or wide) format to a panel (or long) format (Jaquette & Parra, 2014). This means that rather than each institution existing on a different row with a variable for a single year in each column, the data must be converted to a longitudinal format where institutions exist in multiple rows where each row represents a given year, and variables exist one time each in the columns. The *Restructure* function within SPSS was employed in order to transform the dataset into a panel format suitable for longitudinal analysis.

In the longitudinal format, this dataset covered eight periods of time where Year 1 equaled 2008-2009 and Year 8 equaled 2015-2016. Although the variables included in the model from IPEDS had already been appropriately labeled with the academic year in which they were reported, for some of the variables it was necessary to align different academic years together when converting the data from wide to long format. Graduation and retention rates that are reported in a given academic year do not pertain to the same cohort of students as metrics reported for incoming students in that same year. Rather, 6-year graduation rates and 1-year retention rates correspond to the students who entered the institution 6 years or 1 year prior, respectively. Therefore, the variables converted from wide to long format were aligned as follows in Table 5.

Table 5

Cohorts of Data Converted from Wide to Long Format

Year	All other variables	Full-time retention rate	6-Year graduation rate
1	2008-2009	2009-2010	2014-2015
2	2009-2010	2010-2011	2015-2016
3	2010-2011	2011-2012	2016-2017
4	2011-2012	2012-2013	2017-2018
5	2012-2013	2013-2014	2018-2019
6	2013-2014	2014-2015	2019-2020*
7	2014-2015	2015-2016	2020-2021*
8	2015-2016	2016-2017	2021-2022*

Note. Asterisks denote years in the future for which data did not yet exist, but they were added to the dataset in order that there would be 8 years of variables to allow for successful conversion from wide to long format.

The resulting dataset created from a sample of 1,681 institutions contained 13,448 cases spanning 8 years, comprising 166,270 valid observations with 35,450 missing observations. The last 3 years for 6-year graduation rates, noted with asterisks in Table 5, were not available in IPEDS as they stretch too far into the future, but it was necessary to add them to the dataset in wide format so that 8 years of each variable would be available for all variables when converting to long format.

A fixed-effects model considers attributes of the sample that are invariant over time (Blumenstock, 2011) while performing calculations of effect size based on time-variant variables included in the model. In this case, the two fixed-effects in the model were institution and year. One way to account for these effects in the regression is to create dummy variables for each institution and each year in the sample, which one may accomplish using the *Create Dummy*

Variables procedure in SPSS. The outcome of this procedure was that a new binary variable was created for each of the 8 years as well as every institution in the dataset where the value 1 represented the year or institution in that row, and 0 represented all years or institutions not in the row. As a result, several thousand new variables were added to the dataset.

Missing Data

Before any statistical analysis is performed, it is also necessary to identify and account for missing data within a dataset. In statistical modeling, three basic types of missing data exist. The first is missing completely at random (MCAR), which, as the name suggests, means that the missingness of those particular values are random and are not impacted by other variables in the study. The second type is missing at random (MAR) whereby missing values occur because of the effects of other observed variables in the dataset. The last is missing not at random (MNAR). In this situation, data may be missing due not only to their relation with other variables but also due to the value of the missing data point itself (Fielding et al., 2008; Pedersen et al., 2017).

In addition to MNAR, MAR, and MCAR, Young and Johnson (2015) explained that there are two possible types of missing data in longitudinal analyses: within-wave and whole-wave. Within-wave missing data are individual survey question answers or data points that have no value for a particular point in time, whereas whole-wave missingness occurs when an entire set of responses is missing for a record (Young & Johnson, 2015).

One possible way of handling missing values is to exclude all records with any missing data from the study altogether, otherwise known as listwise deletion (Adachi, 2005). However, this could limit the number of institutions in the sample unnecessarily. Instead, using a method of imputation to provide values for the missing data points can result in more valid results since all institutions in the sample become complete and will thus be included in the analysis.

One option for imputing the missing values is to perform mean imputation whereby the mean value for the given variable is calculated and entered as the value for any institution missing a value for that same variable (Adachi, 2005). A second more robust option is to use an imputation algorithm, which computes a random value for each missing data point based upon information included elsewhere in the dataset (Adachi, 2005).

Young and Johnson (2015) performed an experiment whereby they ran fixed-effects and other types of analyses on a dataset with a variety of within-wave and whole-wave missing values. They began with a perfect dataset without any missing values and created several alternative versions of it with different types of missing values. Next the researchers used multiple imputation to enter values for within-wave missing values, for whole-wave missing values, and for both types of missing values. For some of these datasets the missing values were imputed in the long format while others were entered in the wide format. Finally, they compared the results of their statistical analyses of all the various versions of their dataset to the original complete dataset. Young and Johnson (2015) found that for fixed-effects modeling "[w]hen only the observed waves were included in the analysis…data imputed in the long structure produced estimates that were comparable to those imputed in the wide structure" (pp. 289). They also noted that the model in which whole-wave missing values were imputed had the worst results compared to the perfect model.

Because Young and Johnson (2015) noted that most statistical packages have a mechanism to account for whole-wave missingness built into their analyses, and based upon the results of their study, SPSS's internal commands were used to impute the missing values in the dataset used for this study. First, the *Analyze Patterns* function was utilized to identify the variables that contain missing data points and to illustrate how identifiable patterns in

missingness might impact the later data analysis. Next, the *Multiple Imputation* function was employed to enter values for missing data based upon regressions performed by the software.

Upon examining the patterns and amounts of missing data present in the sample at hand, the USNWR peer assessment scores were clearly missing the most values. However, the data that were absent for the USNWR variables was not truly "missing" in the traditional sense. The *U.S. News* publication ranks hundreds of institutions on each of its primary four annual lists, but this by no means encompasses all colleges in the United States. Therefore, while some values were actually missing from the list, many of the institutions without values for these variables were simply not ranked by *U.S. News* and therefore had no value, missing or otherwise. Because of this—the fact that many of these missing values were not truly missing but rather did not exist at all—it was decided that the USNWR variables would not be imputed to calculate missing values.

In 1987, Donald B. Rubin published a seminal book on the topic of multiple imputation entitled *Multiple Imputation for Nonresponse in Surveys.* As a matter of general practice, Rubin recommended performing 3–5 multiple imputations in order to create a dataset worth using for statistical analyses (1987). However, in the past two decades this recommendation has become increasingly challenged as several studies have been published illustrating the need and necessity for more imputations (Bodner, 2008; Spratt et al., 2010; von Hippel, 2020) to generate more consistently replicable results when multiple imputation is repeated on the same original data.

In the case of this study, apart from the USNWR rank data discussed above, the next variable missing the most values was 6-year graduation rate, which was missing 41.4% of its values; further discussion regarding this particular dependent variable is found in the subsequent

section. A summary of missing values in the original dataset spanning 8 academic years can be found in Table 6.

Table 6

Summary of Missing Values

Variable	Percent missing values	Number valid values
USNWR peer assessment score	52.6%	6,374
6-Year graduation rate - degree within 150% of normal time	41.4%	7,886
Avg 25th and 75th percentile SAT Critical Reading score plus Math score	29.6%	9,472
Percent of full-time first-time undergraduate students in low-SES bracket	6.8%	12,538
Percent full-time first-time female applicants	6.5%	12,568
Acceptance rate	6.4%	12,593
Test-Optional	5.8%	12,674
Full-time retention rate	4.7%	12,815
Percent of racial/ethnic minority undergraduate students enrolled	3.2%	13,015
Full-time undergraduate total enrolled	3.1%	13,033
Percent female undergraduates enrolled	3.1%	13,033
Percent undergraduates receiving grant aid	2.9%	13,057

In an attempt to conduct enough multiple imputations on the data to produce a replicable dataset with valid results, I chose to perform 20 imputations using a variety of other variables in the sample as predictors in order to calculate possible values where none existed in the original dataset (Bodner, 2008; Spratt et al., 2010; von Hippel, 2020). The USNWR peer assessment score variable was not imputed due to the nature of its missing values discussed above.

6-Year Graduation Rate Analysis

Over the 8 years of data encompassed in this dataset, 41.4% of values for 6-year graduation rates were missing. However, this is not surprising. Due to the necessity of matching graduation rates 6 years in the future to the other data used in this study, variables for 6-year graduation rates had to be included in the model for the academic years 2014-2015 through 2021-2022. Obviously, since some of these years were in the future and therefore could not be obtained from IPEDS, empty variables (prior to imputation) for 6-year graduation rates for academic years 2019-2020 through 2021-2022 were added to the dataset while it was in wide format in order for there to be eight graduation rate variables to align with the corresponding points in time when converting the dataset into long format. Thus, it is unsurprising that with three completely empty years of data plus other randomly missing values scattered throughout the other years that this variable would have such a high amount of total missing values. However, when examining only the 5 years for which data existed (2014-2015 through 2018-2019), just 6.2% of values were actually missing.

Due to the structure of the data at hand, I determined at this point that it would be most beneficial when analyzing 6-year graduation rates in particular to examine only the years for which data were obtainable from IPEDS, separate from the main analysis spanning 8 years. Returning to the original dataset in wide format and beginning with the same sample of 1,681 institutions, the data were converted from wide to long format again, but this time for just 5 years. The cohorts of corresponding academic years can be found in Table 7.

Table 7

Cohorts of Data Converted from Wide to Long Format in 6-Year Graduation Rate Analysis

Year	All other variables	Full-time retention rate	6-Year graduation rate
1	2008-2009	2009-2010	2014-2015
2	2009-2010	2010-2011	2015-2016
3	2010-2011	2011-2012	2016-2017
4	2011-2012	2012-2013	2017-2018
5	2012-2013	2013-2014	2018-2019

These actions resulted in two separate datasets which would be analyzed using the same procedures, with the primary dataset covering 8 years and providing results on five of the dependent variables in the study and the secondary dataset covering 5 years and providing results on only the 6-year graduation rate dependent variable. A summary of missing values from the secondary dataset spanning 5 academic years can be found in Table 8.

Table 8

Summary of Missing Values, 6-Year Graduation Rate Analysis

Variable	Percent missing values	Number valid values
USNWR peer assessment score	54.2%	3,851
Average 25th and 75th Percentile SAT Critical Reading score plus Math score	29.4%	5,933
Percent full-time first-time female applicants	8.9%	7,656
Percent of full-time first-time undergraduate students in low-SES bracket	8.9%	7,661
Acceptance rate	8.5%	7,692
Test-Optional	7.9%	7,741
6-year graduation rate - degree within 150% of normal time	6.2%	7,886
Full-time retention rate	5.4%	7,953
Full-time undergraduate total enrolled	4.3%	8,044
Percent of racial/ethnic minority undergraduate students enrolled	4.3%	8,044
Percent female undergraduates enrolled	4.3%	8,044
Percent undergraduates receiving grant aid	3.8%	8,088

After the secondary panel dataset was constructed, in the same way that missing data were handled in the primary dataset, 20 multiple imputations were performed on the data using a variety of other variables as predictors to calculate values where they were previously missing. Then, the same steps described in the section above regarding creating dummy variables for year and institution were conducted in an identical fashion on this second dataset. Next, the process outlined in the following data analysis section was performed on both the 8-year and 5-year panel datasets separately.

Data Analysis

As previously mentioned, the statistical model employed in this study was fixed-effects analysis of longitudinal data. Before justifying the appropriateness of using a fixed-effects model in this particular situation, it is worth noting the nature of the data used to answer the set of research questions posed in this study. Institutions of higher education in the United States with various forms of test-optional policies have been adopting them throughout the past several decades but each at different points in time. Organizations with an interest in minimizing the role standardized testing plays in the American educational system, such as FairTest, have documented an increase in the number of schools becoming test-optional in recent years.

Because new institutions join the ranks of test-optional schools each year, the number of colleges in this group is constantly changing. However, tracking an accurate list of test-optional colleges and universities is made even more complicated by the fact that not all institutions that have become test-optional have retained their status as such; some schools have only espoused such a policy for a few years before returning to their former test-requiring status. Therefore, the net number of institutions in the test-optional category is continuously in flux. This study seeks to examine the macro-level impacts of the test-optional movement upon American higher education as a whole and over time. However, this can be difficult to analyze when some schools have had test-optional policies for much longer than others and some that formerly had test-optional policies no longer do. It is not as simple as looking at the differences in key variables between the two groups at certain static points in time.

The fixed-effects model, therefore, was used as the most appropriate way to answer the questions at hand primarily because this model inherently controls for the fact that the observed variables change over time (Torres-Reyna, 2007, slide 9). In other words, it provides an

explanation as to how the included independent variable(s) affects the outcome variable(s) in a given sample over a particular span of time while controlling for time-invariant factors. It also allows for the fact that the institutions included in the test-optional category versus the test-requiring category have changed each year. Given that within the sample used in this analysis institutions have adopted test-optional policies at various points in time, the fixed-effects model of panel data makes the most sense for providing valuable results that examine the impact of test-optional policies on college outcomes.

Additionally, "Panel data allows you to control for variables you cannot observe or measure...or variables that change over time but not across entities" (Torres-Reyna, 2007, slide 3). This fact is appropriate for this particular study since so many possible factors exist, which could influence the outcome, but which cannot necessarily all be accounted for even with the most thorough of data collection.

The formula for the fixed-effects panel data analysis is as follows (Greene, 2008, pp. 183):

$$y_{it} = x_{it}\,\beta + \alpha_i + \varepsilon_{it}$$

In this formula, y_{it} is the intercept (or dependent variable) where i is the institution and t is time. X_i is the independent variable whereas β is its coefficient. Next, α_i represents the variables that are observed, and finally ε_{it} is the error term (Greene, 2008, pp. 183; Torres-Reyna, 2007, slide 10).

More specifically, this model is fit to the current study as follows:

$$y_{it} = \text{Test-Optional Policy}_{it}\,\beta + \text{Year}_1\alpha + \text{Year}_2\alpha + \ldots \text{Year}_7\alpha +$$
$$\text{Institution}_1\alpha + \text{Institution}_2\alpha + \ldots \text{Institution}_{1680}\alpha + \varepsilon_{it}$$

In this equation, as with the general formula above, y_{it} represents each one of the seven dependent variables included in the model, the dichotomous test-optional policy variable is the independent variable, each year and each institution in the dataset constitute the observed variables save one each that act as the reference variables, and ε_{it} remains the error term.

Longitudinal analyses may be conducted using either fixed or random effects models, and the decision regarding which to choose lies within the variables in the model themselves. The Hausman specification test for endogeneity provides researchers with a way of making this determination. In the Hausman test, the null hypothesis is that the regressors in the model are exogenous, and the alternative hypothesis is that the regressors are endogenous. Significant results allowing for the null hypothesis to be rejected indicate that a fixed-effects model is the optimal choice, while failing to reject the null hypothesis points researchers to using a random-effects model instead (Chmelarova, 2007; Daryanto, 2020; Glen, 2017). Therefore, the Hausman test was conducted for each of the six dependent variables in this study producing six separate significant results and therefore confirming the choice to conduct fixed-effects models for this particular set of data.

SPSS does not contain a function for the Hausman specification test, but Dr. Ahmad Daryanto of Lancaster University has developed a macro to be installed in SPSS adding this function, as well as two other tests for endogeneity, to the software (2020). The macro, called EndoS, requires balanced data in order to successfully execute the calculations, but the original dataset used in this study prior to imputation was unbalanced. In order to overcome this to successfully run the EndoS macro, the 8-year and 5-year panel datasets were copied, and mean imputation was used to insert values where none previously existed. Linear interpolation could not be used to complete the data in this instance because in some cases values were missing for

adjacent years; therefore, running a linear interpolation function on the data input some missing values but not all of them. Rather, mean imputation created an 8-year panel dataset with complete data for 13,448 cases and a 5-year panel dataset with complete data for 8,405 cases. The Hausman test could then be successfully executed resulting in the rejection of the null hypothesis in all cases, confirming the use of a fixed-effects method of analysis.

In order to produce usable and meaningful results, six separate linear regressions were computed in SPSS using the *Linear Regression* function: one regression for each dependent variable. In each instance, the key independent variable was the dichotomous variable representing an institution's test-optional policy where 0 represented test-required and 1 represented test-optional. The additional independent variables were other predictors considered to have potential to influence change in the dependent variable as well as all but one of the dummy variables for year and institution; one year and one institution were left out of each regression in order to serve as the reference categories. Using this method, results of six regressions were obtained and their results reviewed. A report of these results as well as a discussion of their implications follow in the next two chapters.

Switchers Versus Non-Switchers

In any dataset used for large-scale analysis, some data will naturally be missing for a variety of reasons. In this case, the number of institutions reporting their test score policy increased from 1,518 in 2008 to 1,670 in 2015. The number of test-optional institutions rose steadily across the 8 years in the sample from 255 in 2008 to 430 in 2015. However, the number of test-required institutions fluctuated across time as shown in Figure 2 found in the explanation of sample selection above.

Fixed-effects analyses examine the impact of a set of independent variables on the dependent variable(s) in question across a certain span of time. In the case of this analysis, the independent variable of most interest was the test score policy of the institutions in the sample. This variable was binary as institutions could either be test-required or test-optional. When the independent variable in a fixed-effects analysis is categorical, the results of the analysis illustrate the impact on the dependent variable of switching from one independent variable category to the other. Therefore, this analysis looked at the amount of change that occurred in a set of dependent variables as a result of changing from a test-required to test-optional policy. Out of 1,681 institutions in the sample, 223 institutions changed their test score policy at some point during the 8 years in the study and are thus defined as "switchers." This number comprised 13% of the total sample.

Some institutions in the study remained test-optional for the entire set of 8 years analyzed, and the same is true for a certain number of schools that were test-required for all 8 years. Of those institutions that changed their policy at some point from 2008–2015, Table 9 shows the breakdown of institutions with the number of years out of the total of 8 during which time they had a test-optional policy.

Table 9

Number of Institutions and Number of Years with Test-Optional Policy

Number of years test-optional	Number of institutions
1	67
2	31
3	25
4	29
5	24
6	17
7	30
Total institutions	**223**

Note. Data from IPEDS (2020).

Limitations and Delimitations

Limitations

Within any empirical study, some limitations naturally exist, which prevent the researcher from stating with absolute certainty that the results of the data analysis are infallible. Since the 2000-2001 academic year, NCES has not accepted IPEDS survey data in hardcopy. Instead, institutions have submitted their responses by either completing electronic survey forms or by uploading data files with their responses in pre-specified formats to load directly into the IPEDS database (email from IPEDS help desk 1/2/18). If any human error occurred by institutional employees at the point of data entry or data file creation, which may have compromised the integrity of the data, it would be unknown to any end user of the data including myself. If such errors did occur, they might affect the validity of the results of this study.

The previous section discussed the difference between switchers and non-switchers and these groups' relationship to the results of the analysis. Should this study have sought to examine differences between test-optional and test-required institutions at a single, specific point in time, the type of analysis performed on the data would have been different, inevitably influencing the results. However, the main intent of this study was to illustrate to institutions considering a test-optional policy how their choice might impact other institutional outcomes of concern; therefore, a longitudinal, fixed-effects method of analysis was employed instead. This decision meant that the 223 switchers in the sample provided the basis for the results of the analysis, which is not a very large number. If a greater number of years had been included in the study, the number of switchers may have been greater, impacting the final results. Similarly, if a different type of analysis comparing test-required and test-optional schools at a single point in time had been used, more institutions in the sample may have formed the basis of the results.

Additionally, if I had restricted institutions in the study to just one type (e.g., public, liberal arts) the results might have been more helpful to institutions of the type included, but this also would have made the sample much smaller. Should the specifications creating the sample have been different, the results of the study would consequently have been different, as well.

In this study, the variable for racial and ethnic diversity rate was calculated using all non-White students and thus included Asian American students. This variable did not represent disadvantaged minority students because Asian American students are often not considered to be underrepresented in educational settings. If Asian American students had not been included in calculating this variable, the results of the analysis could have been applied to disadvantaged minority students specifically as opposed to all minority students, which would have changed the interpretation and recommendations for policy implementation.

The inclusion of rank as published by *U.S. News and World Report* was a decision that was debated for a long time before determining it was relevant enough to the question of change in institutional outcomes to be included in the study. However, the inclusion of USNWR rank data proved to be somewhat of a limitation in and of itself. The data on historical ranking lists are not available electronically and therefore had to be obtained and entered into the dataset by hand. This naturally could have led to some data entry errors. Additionally, USNWR has changed the format of its reporting of rank a great deal over the years including altering the number of institutions and varying the metrics used to calculate rank, which were printed in hardcopies of its annual Best Colleges publications. In most years, four lists were printed but in other years there were just two. The publication has not always printed as many institutions with their ranks in the hardcopy books as on its website, but each year the website's comprehensive lists have been updated; therefore, copies of older more complete lists have been lost. Data on USNWR rank included in this study was as complete as possible, collected by obtaining hardcopies of USNWR's *Best Colleges* books for the years 2008–2015 (excluding 2013, which was never located), but the data would have been more complete had I been able to obtain copies of the full lists printed on USNWR's website in past years. Unfortunately, the publication's unwillingness to share its data with researchers left no choice but to use the lists printed in the *Best Colleges* books, instead.

On a similar note, peer assessment score was used as a proxy for rank reported by USNWR because this measure was printed in the *Best Colleges* books for more institutions than actual rank values and because it is an apples-to-apples comparison across all versions of the *Best Colleges* lists. *U.S. News* usually chooses to publish rank for the top three quarters of each list but the bottom 25% are listed alphabetically with no rank assigned. Peer assessment scores

were generally published for more institutions than rank in the *Best Colleges* books for 2008–2015 and thus were used as a proxy for rank. However, the two are different measurements; therefore, the decision to use peer assessment score as a proxy for rank inevitably impacted the results of the analysis obtained.

Delimitations

Specific decisions made regarding the construction of the sample also helped define the scope of the study. The sample used here was restricted to 4-year degree-granting institutions located in the United States, but it is worth noting that the sample size could have been expanded should these restrictions not been placed upon the group. Institutions outside the U.S. were excluded because the focus here was on the impact of removing SAT and ACT exam scores from the admission process, and these exams are primarily used in the United States. Nevertheless, perhaps the results would have changed if the sample had not been limited to 4-year institutions, or if institutions in other countries with test-optional policies regarding the entrance exams for their respective countries had been added to the sample.

Lastly, I have chosen to use the broadest possible definition of test-optional admissions that is available using IPEDS data because I was less concerned with the differences between test-optional, test-flexible, and test-blind schools and more interested in the differences between test-required institutions and those with any kind of alternative policy in place. However, there is a very large difference between a test-optional policy that only makes standardized test score submission optional for a small portion of applicants who have high school GPAs over a certain threshold and policies that are truly test-blind for all applicants. Should I have decided to restrict my sample to a more narrowly scoped definition of test-optional policies, the impact of more

encompassing test-optional policies upon other institutional outcomes would have been much clearer.

Chapter IV

Results

Introduction

This study seeks to contribute to the existing research regarding the relationship between test-optional admission policies and educational outcomes of postsecondary schools. The intent is to provide empirical evidence to support institutional decisions regarding adopting test-optional admission programs. The research questions explored in this study investigate the impact of test-optional policies on three key areas of higher education: diversity, selectivity, and student outcomes.

This chapter is organized into two parts. It begins with a descriptive analysis of the institutions of the sample in both the 8-year and 5-year datasets used in the study, as the basic characteristics and general profiles of the institutions provide a context for the statistical analyses and allow for an in-depth understanding of the answers to the research questions. Next, the key results of the fixed-effects data analyses will be discussed in response to each of the three research questions, followed by a summary of the chapter.

Descriptive Statistics

Descriptive Analysis of the 8-Year Dataset

This section presents descriptive statistics of continuous variables used in the 8-year dataset. Descriptive statistics including mean, standard deviation, minimum, and maximum for each variable are presented for admission and enrollment variables in Table 10 and for student outcome, selectivity, and rank variables in Table 11.

Table 10

Descriptive Statistics of Admission and Enrollment Variables in the 8-Year Dataset

Variable	*n*	Min.	Max.	*M*	*SD*
Admissions - Full-Time, First-Time Undergraduate Students					
Total applicants	12631	0	92690	5290.61	7964.03
Percent female applicants	12568	0.00%	100.00%	56.12%	0.17
Total admitted	12608	0	36088	2955.03	3965.76
Total enrolled	12614	0	9488	891.41	1198.30
Percent female enrolled	13033	0.00%	100.00%	55.01%	0.17
Avg 25th and 75th Percentile SAT Critical Reading Plus Math score	9470	600	1550	1062.68	136.79
Enrollment - Total Undergraduate Students					
Total full-time enrolled	13033	1	50484	4099.36	5809.62
Total full-time female enrolled	13033	0	25543	2222.56	3031.12
Percent low-socioeconomic status students	12408	0.00%	100.00%	37.46%	0.18
Percent receiving grant aid	13057	0.00%	100.00%	74.73%	18.29
Percent racial/ethnic minority students	13020	0.00%	100.00%	27.25%	0.21

The descriptive analysis in Table 10 shows that for full-time first-time freshmen, the mean of total applicants was 5,290.61 students, the mean of admitted students was 2,955.03 students, and the mean of enrolled students was 891.41 students. These numbers decrease as they move from applicants to admitted students to enrolled students. Women constituted the majority of applicants and enrolled students at institutions in this sample, as the means of these two variables were 56.12% and 55.01%, respectively. Of the institutions in the sample, the average

25th and 75th percentile SAT Critical Reading Plus Math score ranged from 600 to 1550 points (on a scale of 1600) with an overall average of 1062.68 points.

The numbers in the first section of the table above reported by admission offices represent information for full-time first-time students only, but the second section of Table 10 shows information for the entire full-time undergraduate student bodies at institutions in the sample. Across the sample in the 8-year dataset, the average total full-time undergraduate enrollment was 4,099.36 students with a mean full-time female undergraduate enrollment of 2,222.56 students. The sample had an average enrollment of students from low-socioeconomic backgrounds of 37.46%, a mean of 74.73% of undergraduates receiving grant aid, and an average of 27.25% of undergraduate students from racial and ethnic minority backgrounds.

Table 11

Descriptive Statistics of Reputation and Student Outcome Variables in the 8-Year Dataset

Variable	***n***	**Min.**	**Max.**	***M***	***SD***
Acceptance rate	12593	0.00%	100.00%	65.63%	0.20
USNWR peer assessment score	6374	1.30	4.90	2.78	0.54
6-Year graduation rate	7886	0.00%	100.00%	53.74%	19.59
Full-time retention rate	12815	0.00%	100.00%	75.34%	13.33

Table 11 shows the descriptive statistics for four of the dependent variables used in this study related to reputation and student success outcomes. The average acceptance rate, which is the proportion of applicants who are admitted, across all institutions in the sample was 65.63%. In terms of the difference in acceptance rate between the two groups of institutions in the sample, test-required institutions had a mean acceptance rate of 64.5% whereas test-optional institutions had a mean acceptance rate of 69.6%. The most selective institutions in the United States

typically have very low acceptance rates; in fact, in 2019 the average acceptance rate across the 10 Ivy League colleges was just 7% (Pannoni & Kowarski, 2020). Less prestigious institutions must admit a much greater proportion of their applicants in order to achieve a resulting enrolling class that meets their goals. The mean for peer assessment score from the *U.S. News and World Report* annual *Best Colleges* lists, which was used in this study as a proxy for actual rank, was 2.78 on a scale of 5.0.

In terms of student success metrics, 6-year graduation rates and 1-year retention rates were used as dependent variables in this study. Across the 8-year sample, the mean of 6-year graduation rate was 53.74% and the mean of full-time retention rate was 75.34%. For comparison, in 2018 the national averages for 6-year graduation rate and full-time retention rate at four-year institutions in the United States were 59.5% and 81%, respectively (National Center for Education Statistics [NCES], 2020a).

Comparison of Institutional Characteristics of the 8-Year Dataset by Test-Optional Policy

In order to gain an understanding of the differences between institutions in the sample based on their test score policies, various institutional variables were examined to determine if any common characteristics or interesting trends could be identified among test-optional institutions compared to test-required institutions. This section presents descriptive statistics of the categorical variables used in the 8-year dataset broken down by test score policy. Table 12 displays the results of cross tabulation analyses comparing demographic characteristics of institutions with both test-optional and test-required admission policies.

Table 12

Institutional Characteristics of the 8-Year Dataset by Test-Optional Policy

Variable	Test-Required (*n* = 9,917)	Test-Optional (*n* = 2,757)
Control		
Public	3755 (92.5%)	306 (7.5%)
Private Not-for-Profit	6112 (75.3%)	2001 (24.7%)
Private For-Profit	50 (10.0%)	450 (90.0%)
Basic Carnegie Classification 2015		
Baccalaureate	2853 (75.8%)	909 (24.2%)
Master's	4116 (86.5%)	640 (13.5%)
Doctoral	2247 (95.5%)	106 (4.5%)
Special Focus 4-Year	701 (38.9%)	1102 (61.1%)
Region		
Northeast	2441 (67.3%)	1185 (32.7%)
South	3457 (87.0%)	516 (13.0%)
Midwest	2818 (84.7%)	510 (15.3%)
West	1201 (68.7%)	546 (31.3%)
Degree of urbanization		
City	4639 (76.0%)	1467 (24.0%)
Suburb	2348 (74.1%)	819 (25.9%)
Town	2352 (89.2%)	286 (10.8%)
Rural	578 (75.8%)	185 (24.2%)

Control

The control of an institution of higher education refers to whether a school is public, private, or for-profit. In the sample used in this study, institutions that were both test-required

and test-optional were spread among the three groups. While the vast majority of public institutions (92.5%) were test-required, about one quarter of private not-for-profit schools were test-optional. This could relate to the greater flexibility that private schools possess in terms of crafting their own testing policies as opposed to their public institution counterparts who might be more beholden to governmental and state stakeholders. It also could be evidence of the perceived trend that test-optional policies might be more popular among smaller, liberal arts institutions, which are usually private schools rather than large state systems or land grant institutions (Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Syverson, 2007; Zwick, 2007b).

Carnegie Classification

The institutions in the sample were dispersed among the four groups in this categorical variable, with the largest proportion of test-optional institutions in any of the categories (61.1%) being special focus 4-year institutions. The vast majority of institutions in the other three categories were test-required: 75.8% of baccalaureate, 86.5% of master's, and 95.5% of doctoral institutions.

Region

The highest concentration of test-optional schools (n = 1,185) resided in the Northeast region of the United States. The Northeast had more than double as many test-optional institutions as each of the other three regions, and its test-optional schools also comprised the largest proportion of schools in its region (32.7%). However, this was closely followed by test-optional schools in the West, which comprised 31.3% of the total in that region.

Degree of Urbanization

A larger proportion of institutions in all four urbanization level categories were test-required. Almost 90% of institutions in the sample located in town settings were test-required with less than 11% being test-optional. The largest proportion of test-optional institutions were found to be located in suburban locations (26%) with the smallest percent being located in town locations (10.8%).

Descriptive Analysis of the 5-Year Dataset

A descriptive analysis of the 5-year panel dataset used to answer the research question regarding 6-year graduation rates was similar in many ways to those of the 8-year dataset discussed above. This smaller dataset contained 8,405 cases, and the results of this analysis may be found in Table 13. Since the 5-year dataset comprised the majority of data used in the 8-year dataset, the mean values of the variables in each sample did not differ greatly from each other.

Table 13

Descriptive Analysis of Admissions and Enrollment Variables in the 5-Year Dataset

Variable	*n*	Min.	Max.	*M*	*SD*
Admissions- Full-Time, First-Time Undergraduate Students					
Total applicants	7705	0	61556	5030.37	7279.60
Percent female applicants	7656	0.00%	100.00%	56.10%	0.17
Total admitted	7705	0	30382	2813.14	3728.45
Total enrolled	7705	0	9082	891.97	1172.31
Percent female enrolled	8044	0.00%	100.00%	55.20%	0.17
Avg 25th and 75th Percentile SAT Critical Reading plus Math score	5933	620	1525	1065.99	135.03
Enrollment - Total Undergraduate Students					
Total full-time enrolled	8044	1	50484	4082.06	5720.35
Total full-time female enrolled	8044	0	25543	2215.47	2983.97
Percent low-socioeconomic status students	7583	0.00%	100.00%	36.44%	0.18
Percent receiving grant aid	8088	0.00%	100.00%	74.08%	18.65
Percent racial/ethnic minority students	8046	0.00%	100.00%	25.62%	0.21

For full-time first-time students, the mean total applicants in the 5-year dataset was 5,030.37, the mean admitted students was 2,813.14, and the mean enrolled students was 891.97. The mean 25th and 75th percentile SAT Critical Reading plus Math scores were slightly higher in the smaller dataset, at 1065.99 points on a scale of 1600. The mean percent of students enrolled from low-socioeconomic backgrounds was 36.44% while the mean percent of students enrolled of racial/ethnic minorities was 25.62%. The institutions in the 5-year sample had an average of 74.08% of undergraduate students receiving grant aid.

Table 14

Descriptive Analysis of Reputation and Student Outcome Variables in the 5-Year Dataset

Variable	*n*	Min.	Max.	*M*	*SD*
Acceptance rate	7692	0.00%	100.00%	65.30%	0.19
USNWR peer assessment score	3851	1.30	4.90	2.78	0.52
6-Year graduation rate	7886	0.00%	100.00%	53.74%	19.59
Full-time retention rate	7953	0.00%	100.00%	75.11%	13.13

In regards to reputation, the average acceptance rate of this group of schools was 65.30% compared to 65.63% in the larger sample. The peer assessment score reported on the USNWR annual Best Colleges lists had a mean score of 2.78 on a scale of 5.0. In terms of student outcome variables in the 5-year dataset, the 6-year graduation rate average in this analysis was 53.7% and the full-time retention rate average was 75.1%.

Comparison of Institutional Characteristics of the 5-Year Dataset by Test-Optional Policy

As was the case with the 8-year dataset, the 5-year dataset was examined from a test score policy perspective. The results in Table 15 show the descriptive statistics of various categorical variables separated by test-required and test-optional institutions in the sample.

Table 15

Institutional Characteristics of the 5-Year Dataset by Test-Optional Policy

Variable	Test-Required (*n* = 6,179)	Test-Optional (*n* = 1,562)
Control		
Public	2292 (91.9%)	203 (8.1%)
Private not-for-profit	3856 (77.3%)	1134 (22.7%)
Private for-profit	31 (12.1%)	225 (87.9%)
Basic Carnegie Classification 2015		
Baccalaureate	1787 (77.1%)	530 (22.9%)
Master's	2564 (88.2%)	342 (11.8%)
Doctoral	1395 (95.9%)	60 (4.1%)
Special Focus 4-Year	433 (40.7%)	630 (59.3%)
Region		
Northeast	1557 (69.8%)	674 (30.2%)
South	2134 (88.4%)	280 (11.6%)
Midwest	1761 (86.3%)	279 (13.7%)
West	727 (68.8%)	329 (31.2%)
Degree of urbanization		
City	2886 (77.7%)	829 (22.3%)
Suburb	1469 (76.2%)	459 (23.8%)
Town	1468 (89.7%)	168 (10.3%)
Rural	356 (77.1%)	106 (22.9%)

Control

In terms of institutional control, the largest proportion of test-optional schools were private not-for-profit at 87.9% but this constituted just 225 institutions in the sample. Less than 10% of public schools in the sample were test-optional compared to 22.7% of private not-for-

profit schools. However, the most test-optional institutions in the sample were concentrated in the private not-for-profit group, with 1,134 institutions in the sample falling into this category.

Carnegie Classification

In terms of basic Carnegie Classification, most categories of schools had more test-required institutions than test-optional except for the special focus 4-year group, as was the case in the 8-year dataset. Test-optional institutions comprised 59.3% of the special focus 4-year category, whereas test-required institutions comprised 95.9% of doctoral institutions, 88.2% of master's institutions, and 77.1% of baccalaureate institutions in the sample.

Region

As with the 8-year panel, the largest number of test-optional schools (674) in the 5-year dataset resided in the Northeastern region of the United States with the rest scattered more evenly across the remaining three regions. As a percentage of their regions, test-optional institutions constituted 31.2% of schools in the West and 30.2% of schools in the Northeast.

Degree of Urbanization

As a portion of institutions in their group, test-optional schools made up almost the same percentage of institutions in cities (22.3%), suburbs (23.8%), and rural areas (22.9%). However, they comprised just 10.3% of institutions in town locations. The majority of institutions in all four categories of degree of urbanization did not have test-optional policies.

Descriptive Analysis of Switchers Versus Non-Switchers

In addition to the overall descriptive statistics of the entire sample, this section presents the descriptive analysis of the group of institutions by switcher institutions versus non-switchers. The results of this analysis are displayed in Table 16.

Table 16

Mean Variable Values for Switchers vs. Non-Switchers

Variable	Mean: switchers	Mean: non-switchers
Admissions: full-time, first-time undergraduate students		
Total applicants	5511.58	5256.22
Percent female applicants	57.73%	55.87%
Total admitted	3296.90	2901.72
Total enrolled	866.25	895.33
Percent female enrolled	56.96%	54.71%
Avg 25th and 75th Percentile SAT Critical Reading plus Math score	1026.01	1068.06
Enrollment: total undergraduate students		
Total full-time enrolled	4160.68	4089.77
Total full-time female enrolled	2300.13	2210.42
Percent low-socioeconomic status students	38.38%	37.32%
Percent receiving grant aid	74.73%	74.73%
Percent racial/ethnic minority students	28.31%	27.08%
Additional dependent variables		
Acceptance rate	65.20%	65.70%
USNWR Peer Assessment Score	2.70	2.80
6-Year graduation rate	52.26%	53.98%
Full-Time retention rate	74.29%	75.51%

As shown in Table 16, institutions that switched their test score policy at some point during the course of 8 years received more total applicants and more applications from women on average than the non-switcher institutions. Switchers also admitted more students and enrolled more female first-year students but ultimately enrolled fewer total first-year students on average than non-switchers. The average SAT score for switchers was 1026.01 points, which was considerably lower than average for non-switchers with 1068.06 points. This is quite interesting, as it is generally assumed that with fewer students submitting test scores and with only students submitting scores that feel they did well on the exam, average SAT scores would actually rise at test-optional institutions, but the data in this sample show the opposite.

Switcher institutions, on average, had larger total full-time undergraduate populations and enrolled more female undergraduates, more students from low-socioeconomic backgrounds, and more students of racial and ethnic minority backgrounds. The average amount of students receiving grant aid was the same across both subgroups of the sample. The average acceptance rate was half a point higher for non-switchers, and the average USNWR peer assessment score was 0.1 points higher at non-switcher institutions. Non-switchers also had, on average, higher 6-year graduation rates and full-time retention rates than switchers, with each variable averaging approximately 1.5% higher for non-switchers than switchers in this sample.

Fixed-Effects Panel Data Analyses

Research Question 1: Diversity Rates

Research Question 1 in this study was: Are the diversity rates of test-optional institutions significantly different from those of test-requiring institutions, as measured by the proportion of

racial and ethnic minority students and Pell Grant recipients enrolled? The results of the fixed-effects panel data analysis for racial/ethnic minority students enrolled are presented in Table 17.

Table 17

Results of Analysis of Racial/Ethnic Minority Students Enrolled

Variable	**Coefficient**	**Std. Error**
Test-Optional	0.012***	0.000
Full-Time undergraduates enrolled	-2.963E-6***	0.000
Percent female enrolled	0.207***	0.002
Percent receiving grant aid	0.000***	0.000
Total applicants	-2.723E-8	0.000
Average SAT Critical Reading plus Math score	-3.772E-5***	0.000
Cohorts		
2009	0.006***	0.000
2010	0.025***	0.000
2011	0.037***	0.000
2012	0.044***	0.000
2013	0.055***	0.000
2014	0.065***	0.000
2015	0.072***	0.000
Constant	0.824***	0.004
R Square	0.951	
Adjusted *R* Square	0.951	
F	3184.092***	

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

This model produced a statistically significant *F*-test result, indicating that the model itself was valid. The adjusted *R* square was 0.951, indicating that 95.1% of the variance in racial/ethnic diversity in enrollment was explained by factors included in the regression model.

The fixed-effects regression results indicate that there is a statistically significant relationship between a test-optional freshman admission policy and the diversity rate of the student body as measured by racial/ethnic minority students enrolled, while controlling for the other variables in the model. According to these results, on average, the diversity rate as measured by racial and ethnic minority student enrollment increases by 0.012% when an institution adopts a test-optional policy ($\beta = 0.012$, $p < 0.001$) when controlling for the other variables in the model.

The focus of this study is test-optional policies. However, the other independent variables in the model provided significant results in terms of their relationship to racial/ethnic diversity rate with the exception of total applicants. Although some of the coefficients were not of any substantial amount, one must consider the unit of analysis. For example, the results indicate that a one-unit change in full-time total enrollment is associated with a -2.963E-6 unit change in racial/ethnic minority student enrollment. However, a one-unit change in full-time total undergraduate enrollment is only one extra enrolled student. If we multiply each value by 10,000, we can say that an increase of 10,000 full-time undergraduate enrolled students is associated with a 0.03% decrease in racial/ethnic minority enrollment when controlling for the other factors in the model. Although this impact is still small, it puts the effect of the independent variable upon the dependent variable into a more understandable format.

Similarly, a one-unit change in average 25th and 75th percentile SAT Critical Reading plus Math score, which is 10 points, was associated with a very small change in racial/ethnic

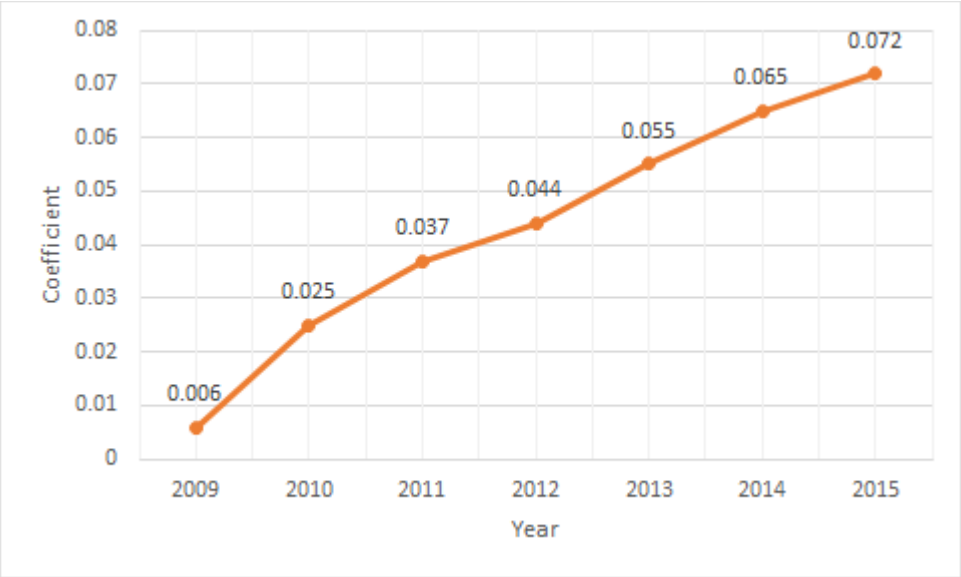
minority enrollment. The change indicated by these results was just -3.772E-5% (β = -3.772E-5, $p < 0.000$), which could also be interpreted as a 1000-point change in average SAT score being associated with a -0.0038% change in racial and ethnic minority student enrollment. More favorable results would have indicated a positive relationship between these two variables rather than an inverse one, but in either case the change is almost negligible.

In contrast, the coefficient produced by the model for percent female enrolled was relatively large compared to the extremely small coefficients for some of the other variables. The results showed that a one percent increase in female enrollment was associated with a 0.207% increase in racial/ethnic minority student enrollment (β = 0.207, $p < 0.000$). This relationship is not directly causal, meaning one cannot state that a 1% increase in female enrollment directly impacts minority student enrollment by 0.21%, but rather the relationship between the two is positively correlated and does seem to be strong. It is plausible that increases in enrollment for both groups of students would be associated with each other, and it is logical that conditions and factors that make an institution more accessible to students from one of these groups would also make it more inviting to students of the other group. Of course, one must also consider the fact that about half of the minority students enrolling will also be women, so crossover between the groups exists, as well.

The results of the time fixed-effects were positive, indicating that compared to the reference year (2008), the other 7 years each had a statistically significant increase in racial/ethnic diversity. In addition, the change in racial/ethnic diversity among institutions as compared to 2008 increased steadily over the 8 years in the model as shown in Figure 3.

Figure 3

Time Fixed-Effects of Racial/Ethnic Minority Diversity Rate Compared to Reference Year (2008)



In addition to race and ethnicity, diversity was examined in this study by using the percent of Pell Grant recipients as a proxy measure for indicating students from low-socioeconomic backgrounds. The results of the analysis of low-SES student enrollment are shown in Table 18.

Table 18

Results of Analysis of Low-Socioeconomic Level Students Enrolled

Variable	Coefficient	Std. Error
Test-Optional	0.025***	0.001
Full-Time undergraduates enrolled	5.580E-6***	0.000
Percent female enrolled	0.078***	0.003
Percent receiving grant aid	0.001***	0.000
Total applicants	-6.878E-6***	0.000
Average SAT Critical Reading plus Math score	0.000***	0.000
Cohorts		
2009	0.058***	0.001
2010	0.101***	0.001
2011	0.097***	0.001
2012	0.093***	0.001
2013	0.094***	0.001
2014	0.097***	0.001
2015	0.093***	0.001
Constant	0.648***	0.008
R Square	0.775	
Adjusted *R* Square	0.774	
F	563.811***	

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

The *F*-value for the low-socioeconomic level enrollment model was significant indicating a valid regression model was used. The overall variance in enrollment of low-SES students that was explained by the model was 77.4%.

The test-optional variable in this analysis had a coefficient of 0.025 ($p < 0.000$), meaning

that schools that choose a test-optional admission program experience, on average, a 0.025% increase in enrollment of students from low-socioeconomic backgrounds when switching from a test-required policy and when controlling for other factors. The results also indicate that a one percent increase in female enrollees is associated with a 0.078% increase in low-SES enrollment ($p < 0.000$). As was the case with the racial/ethnic minority enrollment model, the coefficients for the 7 cohort years were positive, indicating that the time fixed-effects in the model produced significant results for all years as compared to the reference year, 2008.

Research Question 2: Reputation and Selectivity

The second research question in this study sought to answer: Do institutions with test-optional policies gain a competitive advantage in reputation and selectivity as measured by a decrease in acceptance rates and/or an increase in national rankings in the *U.S. News and World Report*? The dependent variables indicated in this question were analyzed separately, and the results of the acceptance rate panel data analysis are presented in Table 19.

Table 19

Results of Acceptance Rate Analysis

Variable	Coefficient	Std. Error
Test-Optional	0.011***	0.001
Full-Time undergraduates enrolled	1.078E-5***	0.000
Percent female enrolled	-0.151***	0.004
Percent receiving grant aid	0.000***	0.000
Total applicants	-1.286E-5***	0.000
Average SAT Critical Reading plus Math score	-6.374E-5***	0.000
Cohorts		
2009	0.006***	0.001
2010	0.001	0.001
2011	-0.004***	0.001
2012	-0.007***	0.001
2013	0.003**	0.001
2014	0.015***	0.001
2015	0.018***	0.001
Constant	0.727***	0.001
R Square	0.734	
Adjusted *R* Square	0.733	
F	451.872***	

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

Table 19 shows that the acceptance rates at schools that leave a test-required policy and become test-optional are 1.1% higher, on average, when controlling for the other independent variables ($\beta = 0.011$, $p < 0.001$). Acceptance rate is perceived as a measure of selectivity with a lower acceptance rate being associated with more prestige since a lower acceptance rate gives the

perception of more difficulty obtaining an offer of admission from the institution. These results indicate that there is a relationship between the two variables, albeit quite modest, but it is the opposite of what might be desired by institutions considering test-optional programs, as the results show an increase in acceptance rate rather than a decrease.

Of the other independent variables included in the model, each produced statistically significant results. For every one percent increase in female enrollees, institutions see a decrease of 0.15% in acceptance rate ($\beta = 0.151$, $p < 0.001$). Additionally, for every 10,000 additional applications, acceptance rate decreases by 0.13%. Unlike the variable for test-optional policies, positive changes in female enrollment and total applicants were associated with negative changes in acceptance rate. Both of these inverse relationships would be more welcome by admission offices seeking ways to decrease their admission rates.

In terms of time fixed-effects of the acceptance rate model, there was a statistically significant change in acceptance rate in each of the cohort years when compared to the reference year, 2008, with the exception of 2010. Four years produced positive coefficients, and two produced negative coefficients. The relationship between the two variables, acceptance rate and time, is not linear in either direction; rather, average acceptance rates for schools included in the study appear to fluctuate across time.

The overall fitness of the acceptance rate model to the research question as explained by the adjusted *R* square value of 0.733 means that 73.3% of the variance in acceptance rate was explained by variables included in the model.

The second measure of reputation included in the model was rank on *U.S. News and World Report's* annual *Best Colleges* lists. As previously stated, peer assessment score was used as a proxy for rank. The results of the USNWR peer assessment score analysis are found in Table

20. The *F*-test for this model was significant, indicating that the model itself was valid. The adjusted *R* square value shows that 95% of the variance in USNWR peer assessment score was explained by the model.

Table 20

Results of USNWR Peer Assessment Score Analysis

Variable	**Coefficient**	**Std. Error**
Test-Optional	0.001	0.002
Full-Time undergraduates enrolled	1.243E-5***	0.000
Percent female enrolled	0.082***	0.016
Percent receiving grant aid	0.000***	0.000
Total applicants	-1.005E-6***	0.000
Average SAT Critical Reading plus Math score	1.127E-5	0.000
Cohorts		
2009	0.017***	0.002
2010	0.050***	0.002
2011	0.061***	0.002
2012	0.059***	0.002
2013		
2014	0.028***	0.002
2015	0.020***	0.002
Constant	-0.013***	0.002
R Square	0.950	
Adjusted *R* Square	0.949	
F	1881.187***	

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

The primary independent variable of interest in this model for test-optional admission policies did not have a statistically significant relationship to USNWR peer assessment score. However, the fixed-effects model shows that every unit change in female full-time freshman enrollment is associated with a 0.082-point change in peer assessment score ($\beta = 0.082$, $p < 0.001$). Given that the peer assessment score is given on a range of 0.0 to 5.0, a change of 0.082 points might make a difference in overall score. In terms of time fixed-effects of the model, each of the 6 years had statistically significant differences in peer assessment score when compared to 2008. No results exist for 2013 because USNWR peer assessment score data for this year were unable to be found and, as discussed in Chapter III, missing values for this variable were not imputed. Since the values for 2013 were nonexistent in the dataset, the year was excluded from the model.

Research Question 3: Student Success Outcomes

The third research question was: What impact, if any, does the adoption of a test-optional admission policy have on student success as measured by first-year retention and 6-year graduation rates? The results of these analyses are found in Tables 21 and 22, respectively. The *F*-values for each model were significant ($p < 0.001$), indicating valid models in both cases.

Table 21

Results of Full-Time First-Year Retention Rate Analysis

Variable	Coefficient	Std. Error
Test-Optional	0.162*	0.067
Full-Time undergraduates enrolled	0.000***	0.000
Percent female enrolled	2.488***	0.275
Percent receiving grant aid	-3.429E-5	0.002
Total applicants	4.573E-5***	0.000
Average SAT Critical Reading plus Math score	0.005***	0.000
Cohorts		
2009	0.231***	0.053
2010	0.229***	0.053
2011	-0.359***	0.054
2012	0.132*	0.054
2013	0.910***	0.054
2014	0.979***	0.054
2015	0.495***	0.055
Constant	57.173***	0.663
R Square	0.724	
Adjusted *R* Square	0.722	
F	428.116***	

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

For first-year students, full-time retention rate produced a statistically significant coefficient related to test-optional policies when controlling for other factors. Schools switching to test-optional admission programs had, on average, a 0.162% increase in full-time retention ($p < 0.05$) while controlling for other variables. Interestingly, the results showed that a one percent

increase in female enrollees was associated with a 2.49% increase in full-time retention rate when controlling for other variables, which is a fairly significant finding ($\beta = 2.488$, $p < 0.001$).

This analysis showed a statistically significant relationship between average SAT score and first-year retention rate, as well. A one-unit change in average 25th and 75th percentile of the SAT Critical Reading plus Math scores was associated with a 0.005% increase in full-time retention rate. Even though it was statistically significant, this change in retention rate was rather minimal. However, a one-unit change on the SAT constitutes a 10-point change, so if each value was multiplied by 10 one could say that a 100-point change in average SAT scores was associated with a 0.05% change in full-time first-time retention rate. This change is still small, but this way of viewing the results puts them into a more understandable context.

The results of the time fixed-effects regression show that full-time retention was significantly different from 2008, the reference year, for all seven time points in the model. Only the year 2011 had a negative coefficient, indicating that the average retention rate that year was lower than the reference year, 2008, whereas all other years with positive coefficients experienced higher full-time first-time retention rates.

Another measure of student success included in this study was 6-year graduation rates, the results of which are shown in Table 22. The data used to conduct this regression analysis were the 5-year panel dataset.

Table 22

Results of 6-Year Graduation Rate Analysis

Variable	Coefficient	Std. Error
Test-Optional	-0.370***	0.093
Full-Time undergraduates enrolled	0.001***	0.000
Percent female enrolled	10.687***	0.418
Percent receiving grant aid	0.035***	0.002
Total applicants	0.000***	0.000
Average SAT Critical Reading plus Math score	0.005***	0.001
Cohorts		
2009	0.322***	0.054
2010	0.447***	0.055
2011	0.440***	0.056
2012	0.471***	0.056
Constant	12.850***	0.918
R Square	0.870	
Adjusted *R* Square	0.869	
F	683.343***	

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

The results in Table 22 show that adopting a test-optional program resulted in an average decrease of 0.37% in 6-year graduation rates when controlling for other variables ($\beta = -0.37$, $p < 0.001$). Institutions obviously seek to graduate as many students as possible from their colleges in a timely fashion. Obtaining results showing that test-optional policies hurt graduation rates rather than help them are not encouraging to colleges considering going test-optional, but these results only speak to the institutions in the sample and the years included in the study.

In addition, all the other independent variables in the model were significantly related to 6-year graduation rates. The results show that a one percent increase in the percent of females enrolled was associated with a 10.69% increase in 6-year graduation rates when controlling for other variables ($\beta = 10.69$, $p < 0.001$). Additionally, a one-unit increase in the percent of undergraduate students receiving grant aid was associated with a 0.035% increase in 6-year graduation rates ($\beta = 0.035$, $p < 0.000$). Finally, a one-unit change in average SAT scores was associated with a 0.005% increase in 6-year graduation rates ($\beta = 0.005$, $p < 0.000$). Each of these results show positive relationships between these predictors and 6-year graduation rates, and indicate areas of policy focus, which might positively influence the number of students who graduate with a bachelor's degree in 6 years.

The time fixed-effects of this model include the 4-year cohort variables compared to the reference group of 2008. All 4 years show statistically significant and positive results as compared to the reference year.

Summary

This study consisted of data from 1,681 institutions across 8 years and analyzed the relationship of test-optional programs to six key dependent variables. Of those six, only *U.S. News* peer assessment score was not found to have a statistically significant relationship to adopting a test-optional policy. The remaining five were significantly related to the primary independent variable, those being acceptance rate, full-time retention rate, 6-year graduation rate, percent of racial/ethnic minorities enrolled, and percent of low-SES students enrolled. It is clear, therefore, that switching to a test-optional program has an effect on educational outcomes deemed important to American colleges and universities.

What follows in Chapter V is a more in-depth discussion of the results of this study, what can be learned from them, implications for policy, and suggestions for future research.

Chapter V

Conclusion

Introduction

The overarching goal of this study was to contribute to the existing literature on test-optional admission policies and, in particular, gain insight into whether or not these programs enhance access to college and increase institutional and student success outcomes. The questions that guided this study were as follows:

1. Are the diversity rates of test-optional institutions significantly different from those of test-requiring institutions, as measured by the proportion of racial and ethnic minority students and Pell Grant recipients enrolled?

2. Do institutions with test-optional policies gain a competitive advantage in reputation and selectivity as measured by a decrease in acceptance rates and/or an increase in national rankings in the *U.S. News and World Report*?

3. What impact, if any, does the adoption of a test-optional admission policy have on student success as measured by first-year retention and 6-year graduation rates?

Summary of Research Design

This study used a fixed-effects panel regression model to analyze data for a sample comprising 1,681 institutions across 8 years of time in order to assess the impact of test-optional admission policies on institutional outcomes in the categories of diversity, selectivity, and student success. Data were collected from the U.S. Department of Education's IPEDS database with the exception of USNWR rank data, which was the only variable collected from outside IPEDS. The sample was made up of 4-year, degree-granting institutions in the United States. Institutions with open admission policies and other outliers were removed from the sample, such as schools located in U.S. territories, which did not report data to IPEDS on their test score policies.

After the sample selection process was completed, the dataset was converted from wide to long format suitable for longitudinal analysis, and multiple imputation was used to account for missing values. The Hausman test for endogeneity was performed to determine whether a fixed-effects or random-effects model was the appropriate choice, and the results indicated that a fixed-effects approach was correct. Six separate linear regressions were conducted on the data, one for each of the dependent variables in the study: racial and ethnic diversity rate, low-socioeconomic enrollment rate, *U.S. News and World Report* rank, acceptance rate, 1-year full-time retention rate, and 6-year graduation rate. The key independent variable used in this study was a binary variable representing an institution's test-optional or test-required status, along with the following additional predictors: full-time undergraduate enrollment, percent female enrollment, total applicants, percent of undergraduates receiving grant aid, and average 25th and 75th percentile SAT Critical Reading plus Math score.

Findings and Discussion

As discussed in Chapter II, the existing body of research supports that small, private, liberal arts institutions are more likely to find a test-optional admission policy appealing (Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Syverson, 2007; Zwick, 2007b). Similarly, the results of the descriptive analysis in this study show that the majority of test-optional schools included in the sample were private not-for-profit institutions located in the Northeast. The reported span of total enrolled undergraduate students ranged from 1 to 50,484 but the mean total undergraduate enrollment was 4,099.36 students, and the median was 1,912 students, indicating that the majority of institutions in this study could be considered small.

Out of the public institutions in the sample, the overwhelming majority (92.5%) were test-required, but a larger proportion of private not-for-profit institutions (24.7%) were test-

optional. The largest proportion of test-optional institutions, 1,185 colleges, were located in the Northeast comprising 32.7% of institutions in the sample from that region. Test-optional institutions also made up a large portion of institutions in the West region at 31.3%. In terms of degree of urbanization, test-optional institutions in the sample comprised approximately a quarter of colleges in city, suburb, and rural locations, but just 11% of institutions in towns.

The sample used in this study had a mean acceptance rate of 65.63%, a mean 6-year graduation rate of 53.74%, and a mean full-time first-year retention rate of 75.34%. In terms of USNWR peer assessment score, which is reported on a 5.0 scale, the mean value across the sample was 2.78 points. Three other variables included in the descriptive analysis were percent of students from low-socioeconomic background, percent of students of minority racial or ethnic groups, and percent of students receiving grant aid. The mean values for these variables were 37.46%, 27.25%, and 74.73%, respectively. Finally, average test scores were included in this study as a variable representing the average 25th and 75th percentile SAT Critical Reading score added together with the average 25th and 75th percentile SAT Math score. The mean value for this variable across the sample was 1062 points over a range of 600 to 1550.

The cohort of institutions in the sample that were considered to be switchers had an average number of total first-year applications of 5,511.58, a 57.73% average of first-year female applicants, a mean of 866.25 total number of first-year enrolled students, and a 56.96% average of first-year female enrolled students. In contrast, the non-switchers in the sample had a mean of total first-year applicants of 5,256.22 students, an average of 55.87% female first-year applicants, a mean of 895.33 total enrolled first-year students, and an average of 54.71% of female first-year enrolled students. The average 25th and 75th percentile SAT Critical Reading plus Math score was higher for the non-switchers at 1068.06 points compared to 1026.01 points

for the switchers. Average total undergraduate enrollment was higher for the switcher institutions versus the non-switchers, at 4,160.68 students compared to 4,089.77 students.

The non-switchers had higher mean acceptance rates, 6-year graduation rates, and full-time first-year retention rates than the switchers, although all these values were within 2% of each other between the two groups. The non-switchers also had a higher average USNWR peer assessment score value of 2.80 compared to 2.70 for the switchers.

The literature on test-optional admissions suggests that test-optional programs may lead to increases in total first-year applications, enrollment of racial/ethnic minority students, enrollment of lower income students, and reported average SAT scores. The descriptive statistics analyzed in this study show that this holds true for applications as well as enrollment of students from low-socioeconomic backgrounds and racial/ethnic minority backgrounds, but not average SAT scores. It is commonly thought that adopting a test-optional program will prompt only those applicants with fairly high test scores to submit them, whereas those with lower scores will choose to omit them, thereby naturally raising the average SAT score reported by a college or university. However, the average SAT score for non-switchers in this study was actually higher than the switchers.

In the review of the existing literature on the topic of test-optional admission policies, several researchers discussed how a test-optional admission program may assist in the overall institutional goal of increasing diversity of students on a given campus (Epstein, 2009; Espenshade & Chung, 2012, p. 190; Rooney & Schaeffer, 1998; Syverson, 2007; Syverson et al., 2018). The concept of "diversity" has been defined in numerous ways. Diversity may be found on a college campus by looking at students with physical or learning disabilities, various religious backgrounds, different sexual orientations, international students, students of non-

traditional age, racial/ethnic minority students, or students of different socioeconomic backgrounds. In the present study, diversity was measured in two ways: racial/ethnic minority and low-socioeconomic students as a proportion of the total undergraduate enrollment. The variable for the percent of undergraduate students receiving Pell Grants was used as a proxy in this study to represent students of low-SES backgrounds. The decision to define diversity in these two ways was based upon prior research by Belasco et al. (2015) and Syverson et al. (2018) who used the same types of diversity measurements in their research. Belasco et al. specifically noted that there has been limited research on the change in racial/ethnic diversity on college campuses as a result of test-optional admission programs.

With regard to the impact of test-optional policies on student diversity, the fixed-effects regression analysis found that adopting a test-optional freshman admission policy was related to a 0.012% increase in enrollment of students from racial and ethnic minority groups and a 0.025% increase in enrollment of students from low-socioeconomic backgrounds. The results suggest that institutions looking for ways to further open college access to students in these two groups may consider a test-optional admission policy as a means of working towards that goal.

As much literature has shown, students of racial/ethnic minority groups often come from underprivileged backgrounds and under-resourced high schools. As a result, they fail to receive the same level of preparation for the SAT and ACT exams as their counterparts from more affluent areas of the country, both in terms of general high school curriculum and specific test preparation resources (Espenshade & Chung, 2012, pp. 177, 181–182; Murray, 2012, p. 72; Robinson & Monks, 2005; Rooney & Schaeffer, 1998; Rosner, 2012, p. 104; Syverson, 2007; Zwick, 2007a). The results are clear: African American and Hispanic/Latino high school students in the United States score lower on average on the ACT and SAT than their White counterparts.

Because of this reality, students in these disadvantaged groups may believe their scores prohibit them from applying to more prestigious colleges. They might also believe that it is not worth their time or application fee money to apply to colleges who publish an average test score higher than their own. SAT and ACT scores have become such an important part of the American college freshman application process that they have actually become a barrier to some low-income and otherwise disadvantaged students, not only prohibiting them from being admitted to certain institutions but in many cases preventing them from even applying. Because the results of this analysis show that test-optional programs are associated with increased enrollment of students in these categories, it suggests that removing test scores from the admission process encourages students in these groups to apply to and eventually enroll in colleges they might not have previously considered accessible to them.

In addition, the results showed that a one percent increase in female student enrollment was associated with a 0.207% increase in racial/ethnic minority student enrollment while controlling for other factors. This was the largest of the statistically significant coefficients produced by this model. It is interesting to consider the crossover between enrollment of racial and ethnic minority students and female students. The results show a positive relationship between enrollment of these two groups, meaning that if one increases so does the other. Both of these types of students are considered minorities by different definitions, but while underrepresented minority students still face many obstacles to college enrollment, more women than men have enrolled in college in the U.S. every year since 1979 (National Center for Education Statistics [NCES], 2019). Perhaps the positive correlation between their enrollment can be explained by the fact that institutions that are actively working to provide welcoming and

supportive environments for racial and ethnic minority students might also seem appealing to female students, and vice versa.

The results of the analysis of the relationship between racial/ethnic diversity rate and average SAT score produced statistically significant results but with a very small coefficient (-3.772E-5, $p < 0.000$). A one-unit change (10 points) in average SAT scores, therefore, has a negligible impact upon racial/ethnic minority enrollment, but of course this is to be expected with just a 10-point change in scores. Even if the unit of analysis is multiplied by 1,000, the change in racial/ethnic minority student enrollment is still very small: -0.038%. Although at first glance these results seem discouraging, they actually support prior research indicating that neither test-optional programs nor increasing the diversity of students on campus will have a substantially negative effect upon the academic quality of the student body (Epstein, 2009; Robinson & Monks, 2005; Syverson, 2007). Average SAT scores are still reported publicly and are included in USNWR's ranking calculations even for institutions with test-optional programs in place, so there is still a need to consider the impact on average scores when institutions are determining whether or not to go test-optional. However, since these results show that average reported test scores only decrease by a very small amount when switching to a test-optional policy, it should be reassuring to institutions making the decision to go test-optional.

With regard to the relationship between a test-optional admission policy and perceived selectivity of an institution, the variables used in this study were admission rate and rank published by *U.S. News and World Report*, with peer assessment score serving as a proxy variable for rank. Results of the analyses indicated there was a statistically significant relationship between test-optional admission programs and acceptance rates, but not between test-optional policies and rank.

For most prospective college students and their families, acceptance rate is a quick and effortless way to compare institutions against one another especially given the fact that websites and services that provide basic information about postsecondary schools (e.g. Peterson's Guide, Niche, Cappex) publish this metric for almost all institutions on their platforms. It is easy to find, easy to digest, and easy to compare across the higher education landscape. Therefore, the ability of a college to employ tactics that might decrease its acceptance rate can be beneficial to the institution. However, the results of this study show that test-optional admission programs are associated with a 1.1% increase in acceptance rate. Although schools might be wary of this, an increase of only 1.1% is quite minimal, and perhaps institutions would be willing to make this small sacrifice in the name of achieving other more important enrollment goals.

These results actually suggest that acceptance rate might not be an appropriate measure for the prestige and reputation of an institution. Prospective students and their families, as well as professionals in the higher education field, look to acceptance rate as a way of gauging how tough it is to be accepted to a college or university, concluding that the more difficult it is to gain admission, the better the school must be. However, this narrative is driven by the idea that the proportion of applicants admitted dictates the quality of an education at the institution. In fact, the quality of an institution can be defined in innumerable ways and indeed manifests itself differently to each person who interacts with a given institution. In support of this argument, *U.S. News* used to include acceptance rate in its calculations of rankings but removed it from the equation in 2018 as it was perceived to be simply boosting the rank of already prestigious schools without providing any true added value to the methodology of calculating rank (Dix, 2018; Strauss, 2018).

Furthermore, the results of the study show that an increase in 10,000 applicants is associated with a decrease of 0.13% in acceptance rate. Finding 10,000 additional applicants is an unimaginable hurdle for a small college yet might be more achievable for a bigger institution enrolling a large number of students. However, the effort it takes to receive 10,000 additional applications might outweigh the benefit of decreasing acceptance rate by less than half a percentage point, even for a larger college or university with a greater number of admissions personnel. These results provide some context to the consideration of going test-optional in an effort to influence other institutional outcomes.

In the United States, institutions covet better spots on *U.S. News's* annual *Best Colleges* lists for much the same reason as acceptance rates are important: Prospective students, and indeed the public more generally, view rank as a measurement of institutional quality and prestige, perhaps more than any other variable (Luca & Smith, 2013; Meredith, 2004). Hence, institutions view an increase in rank as a desirable outcome and welcome the opportunity to positively influence their rank when possible.

In this study, no significant relationship was found to exist between test-optional policies and USNWR rank, but this could be due to a few specific factors. First, the peer assessment score, which is just one part of the calculation of USNWR rank, was used as a proxy for actual rank for two reasons: The rank values themselves were not published for the bottom portion of institutions in most years, and peer assessment scores are calculated on the same scale across all USNWR ranking lists. Rather than using rank values, which span a wide range of values starting at 1 and ending in the hundreds, peer assessment score exists on a scale of 0.0 to 5.0. Therefore, this measurement naturally has a much smaller variance than actual rank. Additionally, peer assessment score is derived from surveys answered by administrators and officials in higher

education across the country and is thus a measurement of the quality of an institution as seen by the rest of the industry. In essence, peer assessment is a subjective value and one that does not seem to change much from year to year. Whereas other measurements that are based on numerical values (e.g., total undergraduate enrollment, average SAT score) might change to a greater degree across 8 years of data, the variation in peer assessment scores over the years in the sample might not have been as great as the variation in actual rank. Thus, the results of the analysis that show no significant relationship between USNWR rank and adopting a test-optional policy may not paint a totally accurate picture of how the two variables relate to one another since peer assessment score might not have been adequate to be used as a proxy variable.

This study also examined the impact of test-optional admissions policies on student outcomes as measured by 6-year graduation rate and first-year full-time student retention rate. Retention rates are critically important for higher education institutions because they are viewed as a referendum from the students' perspective on their first-year experience on a given campus. Many students are unable to continue their enrollment for a second year at a particular institution for financial, medical, or other personal reasons, which are not a reflection upon their attitudes towards the institution itself or the quality of education they have received. However, some students choose to transfer elsewhere or discontinue their course of study at the college at which they spent their first year because they are dissatisfied with some aspect(s) of the institution. Therefore, college administrators actively track the percent of first-year students that return for a second year and provide academic support and social programming for first-year students to create a positive experience and encourage retention of those students.

Given the importance of first-year retention, this study found that adopting a test-optional policy was associated with a 0.16% increase in full-time student rate of retention from the first to

the second year of enrollment. Interestingly, the results of the retention rate analysis also showed that a one percent increase in female enrollees is associated with a 2.49% increase in full-time retention rate when controlling for other variables, which is a fairly significant finding. It is interesting to consider how increasing the enrollment of women relates to greater retention, but it is already known that females outnumber males in college student enrollment so this result supports that previously known fact. According to NCES, the average full-time retention rate across all postsecondary institutions in the United States rose from 71.7% in 2007 to 75.5% in 2018 (National Center for Education Statistics [NCES], 2020a). This metric does not typically change greatly from year to year. Therefore, a change in the range of 0.16% to 2.49%, as this study found, might actually be considered quite significant in the context of the national average over time.

This study also indicated that increasing an institutional average SAT score by 10 points was associated with a 0.005% in retention rate, which is a very modest change, but changing the unit of analysis provides a better understanding of the relationship between the two variables. If each value is multiplied by 10, it shows that increasing the institutional average of the 25th and 75th percentile Critical Reading plus Math score by 100 points is associated with a 0.05% increase in first-year full-time retention rate which is a larger change.

Institutions welcome opportunities to increase both average SAT scores and retention rates, so knowing how the two are related to one another can be beneficial to policy design. However, prior research on the relationship between college entrance exam scores and retention has provided mixed results, with some studies finding no significant relationship between the two and others finding small but statistically significant results (Reason, 2009; Waugh et al., 1994). A study conducted by Levitz et al. showed evidence that higher SAT scores were related

to better retention rates and that, in fact, the higher the average test scores, the lower the student rate of attrition (1999). If indeed a certain level of interaction exists between students who are more academically prepared entering college, as evidenced by higher SAT scores, and students who continue their college educations past the first year, it is not clear from this study that this relationship is directly causal rather than simply a correlation.

Six-year graduation rates have frequently been used as a measure of student success (Cook & Pullaro, 2010). The results indicate that 6-year graduation rates decreased among institutions that switched from test-required to test-optional by 0.37%. This finding is not encouraging to institutions planning to adopt a test-optional policy, as all institutions hope to have as high of graduation rates as possible, but the change is less than half a percentage point, which is quite small. The mission of any institution of higher education includes efforts to support students from the start of their education through the culmination of their degrees. However, the possibility exists that 5 years of data are not enough when conducting longitudinal analyses in order to determine the true magnitude of the relationship between these independent and dependent variables. It should also be noted that in some cases graduation rates are reported at 200% of normal time (8 years) rather than 150% of normal time (6 years), and examining graduation rates in that manner would change the analysis entirely.

As key metrics of measuring institutional performance in the area of student success, retention rate and graduation rate are often analyzed and reported together. Generally speaking, they often are correlated with one another, as students who ultimately graduate from a given institution would be the same group of students who were retained after the first year (Hosch, 2008). However, the results of this study showed an inverse relationship between these variables when framed in the context of admission test score policies; test-optional programs were

associated with increased first-year retention rates but decreased 6-year graduation rates. This finding deserves attention in that it is counterintuitive to the common understanding of the relationship between retention and graduation rates but also suggests that the results of any longitudinal analysis are the function of the variety of institutions in the sample and the years included in the study.

Implications for Policy and Practice

The results of these analyses show that in enacting a test-optional policy, institutions might seek to solicit more applications from racially and ethnically diverse students as well as low-income students. For admissions officials who contend the SAT and ACT are discriminatory towards certain groups of students, this study provides empirical evidence to support much of the literature on how a test-optional policy can be an appropriate response to level the unfair playing field in college admissions intensified by standardized tests. By allowing students of more varied backgrounds the chance to apply without test scores, admission offices are removing this controversial element of the application process. They must then review test-optional applicants using other means that are, in theory, fairer and more indicative of the students' true academic abilities and character. In this way, a more diverse group of students is admitted, and the institution increases the enrollment of students of these particular profiles. The results of this study show a clear relationship between test-optional policies and a higher rate of racial/ethnic and low-SES diversity amongst enrolled student populations, giving credence to institutions looking for this justification in order to move forward with a test-optional admission policy for first-time applicants. This study also indicates to institutions that test-optional programs might positively impact their first-year full-time retention rate efforts. Retention rates are critically

important to institutions to track and maintain, so any measures that may be taken to increase retention are worth considering.

Overall, none of the six analyses conducted in this study produced seriously concerning results in terms of a greatly adverse effect that might occur if an institution adopted a test-optional policy. The overarching goal of this study was to provide colleges with empirical evidence about the impact that a test-optional policy has on other institutional outcomes. Although some results of this study did not indicate a desirable outcome (e.g., decreased 6-year graduation rates when adopting a test-optional policy), none of the results illustrated a grossly negative impact on any of the dependent variables included, allowing institutions that are considering a test-optional program some comfort in making this policy change.

Suggestions for Future Research

The findings of the analysis on low-SES enrollment and racial/ethnic minority enrollment showed that becoming a test-optional institution is related to increased enrollment in both of these categories. These findings mean that more research is needed on the relationship between these two groups in the context of test-optional admissions programs. By using a different, perhaps larger, sample or by examining the data from the student level of analysis, more information can be learned about how test-optional programs might help these two diverse categories of students gain better access to college. It also means that further research needs to be conducted on the relationship between test-optional admissions and increasing college access for diverse groups as defined in other ways. For example, do test-optional admission policies correlate with increased enrollment of students with disabilities, students of non-traditional age, and so on?

The findings of the acceptance rate analysis did not show that switching to a test-optional policy helped decrease acceptance rate, which is a desirable outcome for institutions. It could be the case that had the sample used in this study been larger or the number of years been greater than eight, a larger change in acceptance rate might have been affiliated with going test-optional and possibly could even have indicated an inverse relationship whereby adopting a test-optional policy was associated with a decrease in acceptance rate instead. Further research in this area would be beneficial to better understand the relationship between these two variables.

This study used peer assessment score as a proxy variable for *U.S. News and World Report* rank, but the section on study limitations in Chapter III explained the disadvantages of doing so. Additional research on the relationship between test-optional admissions and rank on USNWR or other national ranking lists using other ways of analyzing rank other than with peer assessment score would be very interesting to pursue.

The results of the 6-year graduation rate analysis indicated that adopting a test-optional program would decrease 6-year graduation rates rather than increase them. However, this study was only able to capture 5 years of data for this analysis so further research in this area using a larger span of time would certainly be beneficial to better understanding the relationship between these two variables. Other researchers might also consider using 8-year graduation rates (200% of normal time) rather than 6 years to see if any difference occurs in the results. Furthermore, since test-optional programs were associated with a decrease in graduation rate in this study, further research is needed to determine the degree to which test-optional programs impact the graduation rates of different subgroups of students. Additionally, these results showed that switching to test-optional assisted institutions in raising their first-year retention rates. An opportunity exists here to look at the impact and extent of this relationship in greater detail.

This study analyzed the impact of adopting a test-optional policy on six key dependent variables related to institutional and student success outcomes. However, this study was not focused on the academic quality of incoming freshman classes at test-optional institutions, which provides another opportunity for future research. In addition, analysis may also be conducted on factors related to financial aid, such as average loan amounts or loan default rates, in order to assess whether test-optional programs are associated with more or less student debt.

If one assumes that the SAT and ACT are discriminatory against certain subgroups of students and that test-optional policies remove this barrier to college, additional research should be conducted on how these students fare once admitted to institutions they may not previously have found attainable. It would be a disservice to already disadvantaged students who are admitted to a selective college or university if institutions do not provide them with the resources and support they need to be successful in college. In addition to retention or graduation rates, other measures of student success should be considered such as first-year or overall college grade point average, job placement, mid-career salary data, etc. Test-optional programs can also be studied in relation to student and academic support programs in student-level analyses.

Overall, using a longitudinal analysis on a greater number of years of data would undoubtedly be beneficial to better understanding how test-optional programs impact other outcomes at a set of institutions. Additionally, future research could certainly be conducted using case study analysis or qualitative analysis in order to further examine the nuances of test-optional programs on specific college campuses. Each institution is unique in so many ways that using quantitative analysis to try to compare them on the same set of factors might not do justice to the variety of differences among institutions in the sample, to which other forms of analysis might be better suited.

Conclusion

The importance of examining test-optional admission policies at the current moment in time cannot be overstated. As this paper is being finished in 2020, the global COVID-19 pandemic is at its height across the world, throwing innumerable aspects of daily life into chaos. Economies and societies worldwide have experienced immense turmoil as public health and elected political officials try to find ways to balance protecting people's health with sustaining the economy, which provides for its citizens' livelihoods. Just one way in which the pandemic has interrupted normal life in the United States is that many of the national test dates for the SAT and ACT have been cancelled in 2020, leading to substantially less access to take the exams. Because the test dates have been cancelled nationwide, this lack of access to take the tests is felt by college-bound students across the spectrum of American life.

As a result of this unprecedented situation, at least 594 American colleges and universities announced test-optional policies from January to October 2020 (National Center for Fair and Open Testing [FairTest], 2020). Some are implementing 2- or 3-year pilot programs to study the effects of such a program while others are openly stating that the exemption is temporary, only applying to fall 2021 applicants who are most greatly affected by the lack of SAT/ACT test dates in 2020. Nevertheless, the pandemic has provided a tremendous boon to the test-optional movement, and there is no doubt that proponents of this concept hope that institutions who at the outset expect their test-optional programs to be temporary might see enough positive impact on other institutional outcomes that they will be persuaded to extend their test-optional initiatives into the future.

In light of the extremely unusual state of affairs in the world in 2020 given the global pandemic, future researchers will have a great deal to look back on in order to assess how test-

optional programs begun during or as a result of the pandemic have affected institutional outcomes, student success, and college access. The increase in test-optional policies being adopted by colleges across the country each year will allow for larger samples in future research and more statistical power to assess the impact test-optional programs have upon other aspects of postsecondary institutions.

The concept of test-optional admissions is not new in and of itself, but as with any idea it takes time to take hold. The global COVID-19 pandemic of 2020 aside, many institutions have been hesitant to adopt a test-optional policy, being unwilling to be among the first to walk down an uncharted road. Yet in the last few decades the studies conducted and the information published about the validity of the SAT and ACT exams and their perceived bias against certain manner of students have swayed more and more institutions to step away from requiring these exam scores in exchange for a more nuanced and wholistic approach to the application review process. The study presented here has attempted to glean some guiding information for institutions questioning whether or not a test-optional freshman admission policy is the right path for them to take. The results indicate that test-optional programs do indeed contribute to opening access to college to students of low-income and diverse racial and ethnic backgrounds. They are also associated with enhanced institutional outcomes, namely increased 1-year full-time retention rates. Only time will tell if test-optional admissions will prove to be the dominant policy among American colleges and universities or not and whether its integration into the admissions landscape will have a lasting and fruitful effect upon key measures of student success as well as leveling the playing field of college access.

References

ACT, Inc. (2020a). *2020-2021 ACT fee waiver eligibility requirements.* https://www.act.org/content/dam/act/unsecured/documents/FeeWaiver.pdf

ACT, Inc. (2020b). *About the ACT test.* http://www.act.org/content/act/en/products-and-services/the-act-educator/the-act-test.html#order-reg-materials

ACT, Inc. (2020c). *Current ACT fees and services*. http://www.act.org/content/act/en/products-and-services/the-act/registration/fees.htm

Adachi, E. (2005). *The performance of missing data treatments for longitudinal data with a time-varying covariate.* [Doctoral dissertation, University of Texas at Austin]. Texas ScholarWorks. https://repositories.lib.utexas.edu/handle/2152/2229

Allen, J., & Sconing, J. (2005). Using ACT Assessment scores to set benchmarks for college readiness. *ACT, Inc.* https://files.eric.ed.gov/fulltext/ED489766.pdf

Allman, M. (2012). Going test-optional: A first year of challenges, surprises, and rewards. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 169-176). Teachers College Press.

Applebome, P. (1997, April 2). Insisting it's nothing, creator says SAT, not S.A.T. *The New York Times*. http://www.nytimes.com/1997/04/02/us/insisting-it-s-nothing-creator-says-sat-not-sat.html

Atkinson, R., & Geiser, S. (2012). Reflections on a century of college admissions tests. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 23-49). Teachers College Press.

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.

Blumenstock, J. (2011). Fixed effects models. *Econ 174/274: Global Poverty and Impact Evaluation (UC Berkeley).* http://www.jblumenstock.com/files/courses/econ174/FEModels.pdf

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(4), 651–675.

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). Predictions of freshman grade-point average from the revised and recentered SAT I: Reasoning test. *ETS Research Report Series*, 2000(1), i–16.

Briggs, D. C. (2001). The effect of admissions test preparation: Evidence from NELS: 88. *Chance, 14*(1), 10–18.

Buchmann, C., Condron, D. J., & Roscigno, V. J. (2010). Shadow education, American style: Test preparation, the SAT and college enrollment. *Social Forces, 89*(2), 435–461.

Card, D., & Rothstein, J. (2007). Racial segregation and the Black–White test score gap. *Journal of Public Economics*, *91*(11), 2158–2184.

Chmelarova, V. (2007). The Hausman test, and some alternatives, with heteroskedastic data. Louisiana State University and Agricultural & Mechanical College. https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=1935&context=gradschool_dissertations

College Board (2020a). *SAT fee waivers.* https://collegereadiness.collegeboard.org/sat/register/fees/fee-waivers

College Board. (2020b). *SAT registration fees*. https://collegereadiness.collegeboard.org/sat/register/fees

College Board. (2020c). *SAT suite of assessments: Higher education*. https://collegereadiness.collegeboard.org/educators/higher-ed

Common Core FAQ. (2014). https://www.npr.org/sections/ed/2014/05/27/307755798/ the-common-core-faq#q1

Common Core State Standards Initiatives. (2020). http://www.corestandards.org/

Cook, B., & Pullaro, N. (2010). College graduation rates: Behind the numbers. American Council on Education. https://vtechworks.lib.vt.edu/bitstream/handle/10919/83996/CollegeGraduationRates.pdf?sequence=1

Cornwell, C., Mustard, D., & Van Parys, J. (2012). The new SAT and acadmic performance at a flagship public university. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 127-136). Teachers College Press.

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA after removing g. *Intelligence*, *36*(6), 719–729.

Daryanto, A. (2020). EndoS: An SPSS macro to assess endogeneity. *The Quantitative Methods for Psychology*, *16*(1), 56–70.

Devine-Eller, A. (2012). Timing matters: Test preparation, race, and grade level. *Sociological Forum*, *27*(2), 458–480.

Dix, W. (2018, September 14). U.S. News changes its college-ranking methodology again; why you shouldn't care. *Forbes.com*. https://www.forbes.com/sites/willarddix/2018/09/14/u-s-news-changes-its-college-ranking-methodology-again-why-you-shouldnt-care/?sh=19d359ce4896

Dixon, C. (2015, December 1). Test optional or test flexible? What you need to know. *Black Enterprise*. https://www.blackenterprise.com/test-optional-or-test-flexible-what-you-need-to-know/

Dixon-Roman, 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).

Epstein, J. P. (2009). Behind the SAT-optional movement: context and controversy. *Journal of College Admission*, (204), 8–19.

Espenshade, T. J., & Chung, C. Y. (2012). Diversity outcomes of test-optional policies. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 177–200). Teachers College Press.

Fielding, S., Fayers, P. M., McDonald, A., McPherson, G., Campbell, M. K., & RECORD Study Group. (2008). Simple imputation methods were inadequate for missing not at random (MNAR) quality of life data. *Health and Quality of Life Outcomes*, *6*(1), 57.

Freedle, R. O. (2003). Correcting the SAT's ethnic and social-class bias: A method for reestimating SAT scores. *Harvard Educational Review*, *73*(1), 1–43.

Geiser, S., & Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes (ED502858) ERIC. https://files.eric.ed.gov/fulltext/ED502858.pdf

Gilroy, M. (2007). Colleges making SAT optional as admissions requirement. *Education Digest: Essential Readings Condensed for Quick Review, 73*(4), 35–39.

Glen, S. (2017). *Hausman test for endogeneity (Hausman specification test)*. Retrieved July 7, 2020, from https://www.statisticshowto.com/hausman-test/

Greene, W. (2008). *Econometric analysis* (6th ed.). Pearson.

Grissmer, D. W. (2000). The continuing use and misuse of SAT scores. *Psychology, Public Policy, and Law*, *6*(1), 223.

Hiss, W. C. (2011). Defining promise: Twenty-five years of optional testing at Bates College, 1984-2009. *Bates College*. http://www.bates.edu/admission/optional-testing/

Hosch, B. J. (2008, November). Institutional and student characteristics that predict graduation and retention rates. In *North East Association for Institutional Research Annual Meeting* (pp. 1-13).

Hunt, S. D., & Morgan, R. M. (1995). The Comparative Advantage Theory of Competition. *Journal of Marketing*, *59*(2), 1–15. https://doi.org/10.1177/002224299505900201

Jaquette, O., & Parra, E. E. (2014). Using IPEDS for panel analyses: Core concepts, data challenges, and empirical applications. In M.B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (Vol. 29, pp. 467–533). Springer. http://ndl.ethernet.edu.et/bitstream/123456789/40804/1/540.Michael%20B.%20Paulsen.pdf#page=478

Jaschik, S. (2012). Sarah Lawrence will again accept admissions tests. *Inside Higher Ed*. https://www.insidehighered.com/quicktakes/2012/05/23/sarah-lawrence-will-again-accept-admissions-tests#sthash.P0Ia8dhH.vubAkQmY.dpbs

Jaschik, S. (2016). Discord on test scores. *Inside Higher Ed.* https://www.insidehighered.com/news/2016/05/16/act-and-college-board-offer-conflicting-views-how-compare-sat-and-act-scores

Kaplan, Inc. (2016). *A brief history of the SAT*. http://www.kaptest.com/blog/ admission-possible/a-brief-history-of-the-sat-from-kaplan-test-prep/

Khan Academy. (2016) *Official SAT practice*. https://www.khanacademy.org/sat

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average* (Vol. 5). The College Board Research. https://files.eric.ed.gov/fulltext/ED563202.pdf

Kobrin, J. L., Sathy, V., & Shaw, E. J. (2007). A historical view of subgroup performance differences on the SAT reasoning test. *The College Board*.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, *31*(1), 3–12.

Levitz, R. S., Noel, L., & Richter, B. J. (1999). Strategic moves for retention success. *New Directions for Higher Education*, *1999*(108), 31–49.

Lewin, T. (2013, August 2). Testing, testing. *The New York Times*. https://www.nytimes.com/2013/08/04/education/edlife/more-students-are-taking-both-the-act-and-sat.html

Luca, M., & Smith, J. (2013). Salience in quality disclosure: Evidence from the US News college rankings. *Journal of Economics & Management Strategy*, *22*(1), 58-77.

Mattern, K. D., Shaw, E. J., & Kobrin, J. L. (2011). An alternative presentation of incremental validity: Discrepant SAT and HSGPA performance. *Educational and Psychological measurement*, *71*(4), 638-662.

Meredith, M. (2004). Why do universities compete in the ratings game? An empirical analysis of the effects of the US News and World Report college rankings. *Research in Higher Education*, *45*(5), 443-461.

Merton, R.K. (1957). *Social theory and social structure*. Free Press.

Micceri, T. (2010). *Assessing the usefulness of SAT and ACT tests in minority admissions* (ED510111). ERIC. https://eric.ed.gov/?id=ED510111

Morgan, H. (2016). *Estimating matriculation with a focus on financial aid and test optional policies: Data from a liberal arts institution in the Northeast*. [Doctoral Dissertation, Seton Hall University]. Seton Hall University dissertations and theses. https://scholarship.shu.edu/dissertations/2164/

Morse, R. & Brooks, E. (2020). How U.S. News calculated the 2021 best colleges rankings. *U.S. News & World Report*. https://www.usnews.com/education/best-colleges/articles/how-us-news-calculated-the-rankings

Murray, C. (2012). Abolishing the SAT. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 69-81). Teachers College Press.

Nankervis, B. (2011). Gender inequities in university admission due to the differential validity of the SAT. *Journal of College Admission*, (213), 24–30.

National Center for Education Statistics. (2017). *Immediate college enrollment rate*. https://nces.ed.gov/programs/coe/indicator_cpa.asp

National Center for Education Statistics. (2018). *About IPEDS*. https://nces.ed.gov/ipeds/ about-ipeds

National Center for Education Statistics. (2019). *Digest of education statistics*. https://nces.ed.gov/programs/digest/d19/tables/dt19_303.10.asp

National Center for Education Statistics (2020a). *Graduation and retention rates: What is the full-time retention rate at postsecondary institutions?*. https://nces.ed.gov/ipeds/TrendGenerator/app/answer/7/32

National Center for Education Statistics (2020b). *IPEDS archived survey materials*. https://nces.ed.gov/ipeds/use-the-data/annual-survey-forms-packages-archived?section=Admissions

National Center for Fair and Open Testing (2018). *About FairTest*. http://www.fairtest.org/about

National Center for Fair and Open Testing. (2020). *1,365+ Accredited, 4-Year Colleges & Universities with ACT/SAT-Optional Testing Policies for Fall, 2022 Admissions.* http://www.fairtest.org/university/optional

Pannoni, A. & Kowarski, I. (2020). *What it takes to get accepted into the Ivy League*. https://www.usnews.com/education/best-colleges/articles/2017-05-25/what-it-takes-to-get-accepted-into-the-ivy-league

Pedersen, A. B., Mikkelsen, E. M., Cronin-Fenton, D., Kristensen, N. R., Pham, T. M., Pedersen, L., & Petersen, I. (2017). Missing data and multiple imputation in clinical epidemiological research. *Clinical epidemiology*, *9*, 157.

Perez, C. (2002). Different tests, same flaws: Examining the SAT I, SAT II, and ACT. *Journal of College Admission*, (177), 20–25.

Powers, D. E. (1993). Coaching for the SAT: A summary of the summaries and an update. *Educational Measurement: Issues and Practice*, *12*(2), 24–30.

Rask, K. & Tiefenthaler, J. (2012). The SAT as a predictor of success at a liberal arts college. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 118-126). Teachers College Press.

Reason, R. D. (2009). Student variables that predict retention: Recent research and new developments. *NASPA Journal*, *46*(3), 482–501.

Robinson, M., & Monks, J. (2005). Making SAT scores optional in selective college admissions: A case study. *Economics of Education Review*, *24*(4), 393–405.

Rooney, C., & Schaeffer, B. (1998). *Test scores do not equal merit: Enhancing equity & excellence in college admissions by deemphasizing SAT and ACT results*. National Center for Fair and Open Testing.

Rosner, J. (2012). The SAT: Quantifying the unfairness behind the bubbles. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 104-117). Teachers College Press.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Schaeffer, R. (2012). Test scores do not equal merit: Deemphisizing standardized tests in college admissions. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 153-168). Teachers College Press.

Shanley, B. J. (2007). Test-optional admission at a liberal arts college: A founding mission affirmed. *Harvard Educational Review*, *77*(4), 429–434.

Spratt, M., Carpenter, J., Sterne, J. A., Carlin, J. B., Heron, J., Henderson, J., & Tilling, K. (2010). Strategies for multiple imputation in longitudinal studies. *American Journal of Epidemiology*, *172*(4), 478–487.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613.

Steele, C. M. (1998). Stereotyping and its threat are real. *American Psychologist, 53*(6), 680–681.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797.

Stewart, P. (2016). Admissions movement leaving ACT/SAT behind. *Diverse Issues in Higher Education*. https://diverseeducation.com/article/87965/

Strauss, V. (2018, September 12). U.S. News changed the way it ranks colleges. it's still

ridiculous. *The Washington Post*. https://www.washingtonpost.com/education/2018/09/12/us-news-changed-way-it-ranks-colleges-its-still-ridiculous/

Syverson, S. (2007). The role of standardized tests in college admissions: Test-optional admissions. *New Directions for Student Services*, (118), 55–70.

Syverson, S., Franks, V. W., & Hiss, W. C. (2018). *Defining access: How test-optional works.* National Association for College Admission Counseling. https://www.nacacnet.org/news--publications/Research/Defining-Access/

Thomas, M. K. (2004). Seeking every advantage: The phenomenon of taking both the SAT and ACT. *Economics of Education Review*, *23*(2), 203–208.

Torres-Reyna, O. (2007, December). *Panel data analysis: Fixed and random effects using Stata (v. 4.2)* [PowerPoint slides]. Princeton University. https://dss.princeton.edu/training/

U.S. Census Bureau. (2015). *Geographic terms and concepts - census divisions and census regions*. https://www.census.gov/geo/reference/gtc/gtc_census_divreg.html

Von Hippel, P. T. (2020). How many imputations do you need? A two-stage calculation using a quadratic rule. *Sociological Methods & Research*, *49*(3), 699–718.

Walpole, M., McDonough, P. M., Bauer, C. J., Gibson, C., Kanyi, K., & Toliver, R. (2005). This test is unfair: Urban African American and Latino high school students' perceptions of standardized college admission tests. *Urban Education*, *40*(3), 321–349.

Waugh, G., Micceri, T., & Takalkar, P. (1994). *Using ethnicity, SAT/ACT Scores, and high school GPA to predict retention and graduation rates* (ED453704). ERIC. https://eric.ed.gov/?id=ED453704

Wonnell, T., Rothstein, C., & Latting, J. (2012). Predictors of academic success at a highly selective private research university. In J. Soares (Ed.), *SAT wars: The case for test-optional admissions* (pp. 137-150). Teachers College Press.

Young, R. & Johnson, D. (2015). Handling missing values in longitudinal panel data with multiple imputation. *Journal of Marriage and Family*, *77*(1), 277-294.

Zwick, R. (2007a). College admission testing. *National Association for College Admission Counseling*, 1–44.

Zwick, R. (2007b). College admissions in twenty-first century America: The role of grades, tests, and games of chance. *Harvard Educational Review*, *77*(4), 419–428.