



# GRE-Optional Policies in PhD Admissions: Impacts on Student Applications, Yield, and Structural Diversity

Heeyun Kim 

University of Houston

Michael Bastedo 

John Gonzalez

Allyson Flaster 

Yiping Bai

University of Michigan

*Test-optional policies in graduate admissions have been increasingly common, yet their impact on structural diversity remains a topic of debate. To scrutinize the policies' effects on applications, yield, and structural diversity, we utilized 182,638 admissions records across 117 PhD programs at a flagship U.S. public university. Employing heterogeneity-robust difference-in-differences analyses, we found that the adoption of Graduate Record Examinations (GRE)-optional policies increased program selectivity, evidenced by a decrease in acceptance rates. Changes in applicant, admitted, and enrollee demographics were, for the most part, not significant. Our study provides the first empirical insights into the relationship between GRE-optional policies and enrollment outcomes, while also noting the policies' mostly negligible impact on structural diversity in PhD programs.*

**Keywords:** college admissions, test-optional, standardized tests, PhD admissions, structural diversity, doctoral programs

TEST-OPTIONAL policies have been growing steadily over the years and for selective college admissions are now the predominant new policy in the post-pandemic era. The test-optional policy movement peaked in the mid-2000s among selective liberal arts colleges, which questioned the predictive validity of standardized tests in assessing student aptitude (Epstein, 2009). Since then, a growing number of schools have opted out of requiring standardized test scores in college applications, calling for more holistic and individualized approaches to the admissions process (Furuta, 2017). The COVID-19 pandemic further fueled this movement as it reduced applicants' access to standardized testing (Schultz &

Backstrom, 2021). A majority of schools have extended their undergraduate test-optional policies during the pandemic, and many schools are committing to permanently implementing either test-optional or test-free application policies post-pandemic (Elias, 2022).<sup>1</sup> However, at the undergraduate level, there have been notable recent exceptions to the test-optional trend among primarily elite institutions, including Dartmouth, Harvard, Georgetown, and MIT (Friedman et al., 2024).

The test-optional movement has also influenced graduate admissions, often labeled #GRExit on social media. Opposition to the Graduate Record Examinations (GRE) is

growing rapidly, as most programs at leading research-intensive institutions dropped test score requirements during the pandemic (Kell et al., 2021). At the University of Michigan, where this study takes place, the limited availability of GRE testing during the pandemic forced programs to consider a moratorium on the use of the test, which culminated in an official change to the admissions policy, discontinuing the use of the GRE test for any admission decisions effective Fall 2022. Proponents of test-optional policies argue that GRE test requirements disproportionately harm applicants from underrepresented backgrounds associated with race, ethnicity, gender, first-generation, and socioeconomic status (Bleske-Rechek & Browne, 2014; Feldon et al., 2024). In contrast, some raise concerns about the unintended consequences of test-optional policy adoption without sufficient understanding and preparation for holistic review. They argue that removing the GRE may induce admissions committees to consider other biased selection factors, such as the prestige of undergraduate institutions, research experience, and inflated recommendation letters (Kuncel et al., 2001; Nicklin & Roch, 2008; Woo et al., 2023). Furthermore, programs with a large number of international students may feel more uncertain about how reliably a student's academic abilities are captured in grades and courses, given the lack of contextual information and the specifics of their curricula and evaluation metrics (Furuta, 2017; Posselt, 2016).

Previous studies have explored the impact of test-optional policies on student diversity, and the findings to date are mixed, perhaps because few of them have generalizable samples. Studies focusing on undergraduate admissions found limited impacts of test-optional policies on student diversity, while reporting their positive association with related outcomes such as institutional revenue, application count, and numeric ranking (Belasco et al., 2015; Conlin & Dickert-Conlin, 2017; Rosinger et al., 2021; Rubin & González Canché, 2019; Saboe & Terrizzi, 2019). A notable exception is Bennett (2022), who reported modest increases in Pell Grant recipients, racially minoritized students, and female students at private institutions. The findings of studies at the graduate level are also mixed and mostly based on admission to a single major at one institution,

primarily in health disciplines (Cahn, 2015; Dang et al., 2021; Hooker et al., 2022; Langin, 2022; Miller et al., 2019).

Notwithstanding the previous findings about the association between test-optional policies and student body diversity, it is difficult to extrapolate from them to the recent context of doctoral admissions. Unlike at the undergraduate level, doctoral admissions are highly decentralized, with disciplinary logics determining faculty evaluation of applicants based on varying admissions criteria and practices across disciplines (Griffin & Muñiz, 2015; Kent & McCarthy, 2016; Posselt, 2015, 2016). Furthermore, prospective doctoral students differ from college applicants because of their high academic aspirations, previous experiences with the college choice process, and age, all of which influence the different factors they may weigh when considering application and enrollment (Benshoff et al., 2015; Bersola et al., 2014; Xu, 2014; Zhang, 2005).

Given the mixed findings and the paucity of studies encompassing a large population of graduate students spanning multiple disciplines, the relationship between GRE-optional policies and structural diversity requires more empirical investigation to provide policy-relevant insights. Our study aims to answer the following research questions:

- 1) Does the adoption of GRE-optional policies by PhD programs affect applications, selectivity, and yield in subsequent years?
- 2) Compared to PhD programs that require GRE scores, do GRE-optional programs enroll more students from underrepresented backgrounds after discontinuing their GRE requirement?

Utilizing 182,638 PhD admissions records spanning 2009 to 2020 at the University of Michigan, our study leverages a heterogeneity-robust difference-in-differences (DiD) approach to investigate the causal effects of GRE-optional policies across a wide range of disciplines. Our findings indicate that the adoption of GRE-optional policies enhances institutional selectivity in the first year after policy implementation but has no positive impact on structural diversity, particularly among admitted applicants and

enrollees. Although more female applicants were recruited immediately following the policy change, this did not translate into a corresponding increase in the number of admitted and enrolled female students. In addition, the proportion of racially minoritized students decreased among admitted applicants in the first year after implementation. Our study makes significant contributions to the understanding of test-optional policies and graduate admissions. Previous research has focused on single disciplines and often examined only demographic changes among enrolling students, leaving gaps in understanding what occurs throughout the admissions process. Analyzing how test-optional policies influence both student and institutional behaviors, which in turn affect yield and the overall admissions process, is crucial. Our research provides valuable policy implications by identifying strategic interventions to promote equitable admission outcomes in holistic admissions practices.

### Prior Research

The use of the GRE in admissions has long been debated due to concerns about disparities in test scores among different subgroups within the student population. Previous studies found that women, racially underrepresented minorities, and international students were less likely to submit GRE scores and scored lower than their counterparts (Bleske-Rechek & Browne, 2014; Kell et al., 2021). Concerned that GRE scores could be potential gatekeepers in graduate admissions, many graduate programs have made submitting GRE scores optional for applicants (Woo et al., 2023). The testing constraints caused by the pandemic fueled the GRE-optional movement, and many programs have extended these policies. In 2021, there were 37% fewer GRE test-takers than in 2017 (Saul, 2023). The GRE is also comparatively expensive at \$220 per administration.

Despite the wide diffusion of these policies, there is a scarcity of research examining the effects of GRE-optional policies on structural diversity in doctoral admissions. Many previous studies have shown mixed findings, based on a limited range of samples specific to a few disciplines. Focusing on law schools' attempt to revise

standardized test requirements, Rosinger et al. (2022) found that the adoption of GRE-accepting policies, which allowed the submission of GRE scores in place of LSAT, actually decreased racial diversity while increasing the volume of applications and lowering acceptance rates. Utilizing the interview responses of faculty members and admission officers at 30 health programs that went GRE-optional, Cahn (2015) found that such a policy change, as perceived by interviewees, is unlikely to yield consistent positive effects on campus diversity unless paired with other approaches, including proactive outreach. In a randomized study conducted by Dang et al. (2021) in the Department of Epidemiology and Biostatistics, researchers found no difference in faculty evaluations when the admissions committee was given applications with and without GRE scores. In contrast, Posselt et al. (2023) found that the establishment of new admissions routines at multiple levels, initiated by the elimination of GRE score requirements, could interrupt institutionalized inequalities in STEM PhD programs.

Most quantitative studies conducted at the undergraduate level report null impacts of test-optional policies on the composition of the student body. Employing quasi-experimental techniques, including DiD, these studies have found that the adoption of test-optional policies did not significantly change the composition of the student body with respect to gender, socioeconomic status, or race (Belasco et al., 2015; Conlin & Dickert-Conlin, 2017; Rosinger et al., 2021, 2024; Rubin & González Canché, 2019; Saboe & Terrizzi, 2019). An exception was a recent study by Bennett (2022), which found a modest impact of test-optional policies on the increase in Pell Grant recipients, underrepresented, racially/ethnically minoritized (URM) students, and female students at private institutions. Despite standardized tests being the biggest predictor of socioeconomic and racial inequality (Bastedo & Jaquette, 2011; Posselt et al., 2012), prior studies suggest that shifting away from testing requirements has not led to meaningful changes in undergraduate admissions. Even with the change in test requirements, institutions are still able to infer test scores with a great deal of other information on which to base their acceptance decisions (Conlin & Dickert-Conlin, 2017). Further, using other admissions

criteria to make admissions decisions, such as GPA, class rank, and extracurricular activities, may attract applications from traditional college-going students rather than recruiting a diversified student body (Rubin & González Canché, 2019). In terms of applications and yield, test-optional policies have been found to have positive effects by increasing the volume of applications or overall admissions yield (Belasco et al., 2015; Paris et al., 2022; Rosinger et al., 2024; Saboe & Terrizzi, 2019).

Although these previous quantitative studies lay the groundwork for understanding the influence of test-optional admissions on student enrollment demographics, their findings have limited implications for doctoral admissions with such policies. Unlike at the undergraduate level, doctoral admissions are highly decentralized and dominated by faculty decision-makers. Each program constructs an admissions committee with mostly faculty members and (sometimes) doctoral students, whereas undergraduate admissions are administered by professional staff in a central administrative unit. Faculty evaluations of applicants vary by programs or departments where disciplinary logics shape admissions criteria and practices (Posselt, 2015). For instance, economics programs prioritize GRE scores and academic preparation in mathematics due to their foundation in quantifiable evidence and rational thinking, whereas philosophy programs, rooted in interpretivism, emphasize qualitative factors such as the writing sample, letters of recommendation, and personal statements (Posselt, 2016). Further, perceptions about merit and competency in holistic review also vary between faculty and admissions officers, as faculty members often lack the training and hands-on experience required for holistic and contextualized review. Thus, faculty often determine the results of policy implementation.

### **Conceptual Framework**

We utilize theories of college choice and risk aversion, in conjunction with prior research, to understand the behaviors of two key stakeholders involved in test-optional policies in doctoral admissions: students and doctoral programs. The admissions process involves multiple interdependent stages in which students and institutions make

decisions about whether and which to choose among a set of options. During the first stage, students identify a select number of institutions to apply to by acquiring information about those they are considering and taking standardized tests that are required by some institutions within their choice sets (DesJardins et al., 2006). During the subsequent stage, institutions admit students who are likely to be successful, while attending to enrollment goals pertinent to structural diversity, excellence, and revenue generation (Bastedo et al., 2022; Cheslock & Kroc, 2012). In doctoral admissions, where applicants are reviewed by each program, professors play a role as gatekeepers to the discipline, and the way they operationalize merit and diversity plays a significant role in crafting a cohort in that discipline (Posselt, 2014). In the final stage, students decide whether to accept the program's offer, considering a wide range of factors, among which racially minoritized students place significantly more importance on faculty, student, and community diversity, and cost of living (Bersola et al., 2014).

For GRE-optional policies to impact structural diversity, both students' and institutions' behaviors would need to be affected by the adoption of the new policy. There are multiple potential channels by which the adoption of test-optional policies could affect students' application and enrollment decisions. Three likely channels are publicity, application systems, and signaling (Bennett, 2022). Test-optional programs or institutions often enjoy increased publicity through media attention, which, when combined with the allure of applying without test scores, can make these institutions more attractive to students. If increased program and institutional visibility is a focal mechanism behind the impact of test-optional policies on doctoral enrollments, we would expect to see all of the policy impacts occur at the application stage. We would not expect increased publicity to influence yield rates (the proportion of admitted applicants who accept the admissions offer), however, as the factors driving students' decisions about which graduate program to attend center most strongly on programs' academic characteristics, location, and financial support, not the visibility of programs (Bersola et al., 2014; Kallio, 1995; Pooch & Love, 2001).

One of the main rationales for implementing test-optional and test-free policies is a belief that the value of test scores in predicting doctoral success is low relative to the burden that testing puts on potential applicants. Removing the financial and behavioral barriers associated with standardized testing is hypothesized to particularly benefit students from marginalized backgrounds (Millar, 2020; Pallais, 2015). Since requiring test scores as part of the application process may place a larger burden on marginalized students, we would expect application volume to grow more for marginalized students than for non-marginalized students.

Moreover, the modification in application requirements could prompt students to reassess their alignment with institutions, perceiving test-optional policies as signals of valuing individuality and holistic review based on student personhood (Bennett, 2022). If signaling is effective in delivering the message that the programs value diversity, minoritized students are likely to view the programs more favorably at both the application and the enrollment stage (Bersola et al., 2014). Thus, this is the only channel by which we hypothesize that yield may be affected by the policy change.

In contrast, potential changes in institutional behavior resulting from the adoption of test-optional policies might be limited. We use the conceptualization of March's (1994) organizational risk-taking to understand why the size and structural diversity of the admitted applicant pool may not change in response to removing testing requirements. Under this framework, organizations can have varying levels of risk-aversion depending on their risk estimation and risk-taking propensity. Risk estimates are influenced by the level of knowledge decision-makers have. Incomplete knowledge is a threat to a valid estimate of risk, and decision-makers have a tendency to try to reduce uncertainty. When uncertainty is conceptualized as a lack of information, decision-makers extrapolate from available information based on assumption-based reasoning, employing a mental model of the situation based on common beliefs (Lipshitz & Strauss, 1997). While it is useful in making decisions quickly when faced with uncertainty, the utilization of such heuristics can potentially lead to cognitive biases (Tversky & Kahneman, 1974).

In light of incomplete knowledge about applicants, selective doctoral programs are hypothesized to become more risk-averse following the adoption of test-optional policies. The removal of standardized test scores may make faculty more cautious in admissions, as they lose a piece of information they were accustomed to relying upon, leading them to lean more heavily on applicants' conventional achievements, such as undergraduate GPA (UGPA) and undergraduate prestige, to avoid adverse selection. In addition, faculty may lean on their personal and collective professional experiences, which can outweigh other factors during application evaluations (Hernández-Colón et al., 2021). Cognitive biases may further constrain evaluators, potentially causing them to overlook candidates' potential in the context of their undergraduate and life opportunities (Bastedo & Bowman, 2017).

Faculty members may also exhibit risk aversion in student selection. Risk propensity is influenced by expected decision outcomes and the organization's position, and decision-makers are likely to be risk-averse when the expected returns are low and the perceived risks are high (March, 1994). Organizations that are successful and above a performance target are prone to avoid actions that could potentially place them below it (March & Shapira, 1987). Similarly, selective doctoral programs are risk-averse in their admissions practices as they prioritize program prestige and view risk avoidance as economically advantageous (Posselt, 2015). The intersection of prestige and risk aversion may also justify their relative risk aversion, despite potential negative consequences for diversity. The perceived gain from admitting doctoral students from nontraditional backgrounds may be perceived as "gambling" due to the potential financial and reputational impacts of student failure (Posselt, 2014). To faculty, the greatest fear in doctoral admissions is student attrition, as their academic activities and financial resources are directly affected by the failure to recruit students who are a strong academic fit for them and their program (Wall Bortz et al., 2020). Doctoral students are integral to a department's research and teaching activities, which maintain program status.

Although there is some evidence that reducing the information available to decision-makers has the potential to increase discrimination (Agan

& Starr, 2018), there are also a number of potential countervailing forces at work within a university that would mitigate the negative effects of risk aversion on admitting a holistically qualified cohort. In particular, robust PhD fellowship programs and holistic admissions practices can make test-optional or test-free admissions processes more equitable (Posselt et al., 2023). For instance, university-level initiatives, including Minority-Serving Institutions Partnership or Summer Research Opportunity Program, can encourage faculty to broaden their views about merit in admissions. When faculty's risk aversion in evaluating doctoral applicants is considered alongside countervailing forces to broaden applicant pools, we anticipate that the structural diversity of the admitted applicant pool will remain unchanged in test-optional programs following changes in testing requirements.

## Methods

### *Data*

We utilized admissions records from the University of Michigan (U-M). U-M is a Research 1 institution that received more than \$2 billion in sponsored research funds in FY2024 and fields over 100 doctoral programs in the humanities, social sciences, physical sciences, biological sciences, engineering, and a number of multidisciplinary endeavors (Office of the Vice President for Research, University of Michigan, 2024). The university ranks in the top five in the United States in granting doctoral degrees (National Center for Science and Engineering Statistics, 2024), enrolls nearly 6,000 doctoral students, and enrolls more than 1,000 new doctoral students every year.

Our analysis employed administrative data on applications submitted to PhD programs across all disciplines over 12 admissions cycles (2009–2020). Specifically, this includes students who applied for admission between the 2008–2009 and 2019–2020 admissions cycles (e.g., applying in December 2008 to begin studies in Fall 2009) and who entered programs between the 2009–2010 and 2020–2021 academic years. The application and admission decisions for the 2019–2020 admissions cycle occurred prior to

the COVID-19 pandemic outbreak. The data comprises applicants' sociodemographic information, academic background, test scores, and previous undergraduate and graduate institutions, along with admission results and matriculation status. The student-level dataset has 182,638 observations within the analytical timeframe. For our analyses, we aggregated student-level data to the program level, yielding an unbalanced panel of 12,763 observations across 117 programs.

To contextualize the nature of the GRE-optional movement at U-M, since the early 2010s, doctoral programs have exercised flexibility in both formally and informally requiring GRE scores for admission. Because of both the decentralized nature of admissions in doctoral programs (programs individually make decisions about all admissions aspects) and the national movement toward adopting a more holistic admissions process (Furuta, 2017; Woo et al., 2023), U-M doctoral programs at various times have made their own decisions about how much to emphasize the importance of GRE scores. This movement created an initial set of early programs that de-emphasized the use of the GRE for doctoral admissions.

The adoption of GRE-optional policies varied by discipline over time. Programs in Humanities and the Arts led early adoption, beginning in 2010 and steadily growing to 13 out of 23 programs by 2020. Biological and Health Sciences saw rapid adoption starting in 2016, reaching 17 out of 28 programs by 2020. In contrast, Physical Sciences and Engineering and Social Sciences showed minimal adoption, with only a few programs adopting policies during the pre-COVID timeframe. These trends highlight both the disciplinary differences in adoption timing and the acceleration of adoption in the late 2010s.

### Variables

To identify treatment status, we first collected data on test-optional policies by visiting the historically published webpages of each program using the Wayback Machine. It is a digital archive administered by the Internet Archive, which enables users to access historical website data. We collected relevant admissions

information about each program for admissions cycles from 2009 to 2020 and coded them for test-optional status. We next triangulated this information by plotting GRE score submission rates over time for each program and comparing years with a sharp dip in the rate with the website record. We also examined whether any admissions were offered to applicants who did not submit GRE scores prior to the formal adoption of a GRE-optional policy. In rare cases, some programs admitted students without GRE scores immediately before the policy was adopted. In these cases, we relied on the Wayback Machine record to identify the policy adoption year since it was publicly available information, thereby shaping students' application behavior and program considerations. The final dataset contains the years programs shifted from test-requiring to non-requiring for application to each PhD program.

For the treatment variable, we used a dummy variable indicating whether a program required the general GRE test score in a given year ( $0 = No$ ,  $1 = Yes$ ). The treatment group consists of programs that first adopted a GRE-optional policy in the given year, and the control group comprises programs (never-adopters) that had not implemented a test-optional policy as of 2020. The number of programs in each year that required the GRE ("requiring programs") or made the GRE optional ("non-requiring programs"; Table 1).

We constructed three variables from admissions records to address the outcome variables needed to answer the first research question: annual application growth rate (logged application count), acceptance rate (admissions offers divided by applications), and yield rate (enrollments divided by admissions offers). To answer the second research question, we calculated the share of female students; domestic URM students (defined as Black, Latine, Pacific-Islander, and Native American); international students; first-generation college students; Pell Grant recipients; and average UGPA among applicants, admitted applicants, and enrollees. The descriptive summary of the variables used for the analysis can be found in Supplemental Appendix Table 1 (available in the online version of this article).

To estimate the effect of GRE-optional policies on the outcomes of interest, we employed a DiD method. This approach estimates the treatment effect by comparing the outcomes of the treatment group (programs that adopted a GRE-optional policy in a given year) and the control group (programs that had not implemented a test-optional policy as of 2020) before and after policy adoption. Our dataset's panel structure allowed us to apply two-way fixed effects (TWFE) with program and year fixed effects as the baseline model. However, the TWFE's validity may be compromised by heterogeneity in treatment timing and varying effects over time. Goodman-Bacon (2021) highlights that when treatment adoption is staggered, earlier-treated groups may be incorrectly used as controls for later-treated groups, potentially biasing estimates with negative weights. Recent econometric studies suggest alternative estimators to address these issues and improve the validity of the findings in staggered designs (Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfœuille, 2020; Sun & Abraham, 2021).

We leverage the heterogeneity-robust DiD estimator by Callaway & Sant'Anna (2021), which computes group-time average treatment effects. This estimator has several benefits applicable to our study. It addresses the negative weight issue by estimating the average treatment effect for the treated (ATT) for each group and each time period, and it aggregates the group-time ATT using proper weights. Moreover, it imposes a weaker assumption than other DiD estimators would do, assuming the parallel trends to hold after conditioning on covariates  $X$ . It also offers various ways to aggregate group-time average treatment effects for straightforward interpretation of treatment effect within the policy context. Formally, the group-time average treatment effect is defined as

$$ATT(g, t) = E[Y_t - Y_{g-1} | X, G = g] - E[Y_t - Y_{g-1} | X, C = 1], \text{ for } t \geq g, \quad (1)$$

where it is the ATT for the units in the same group  $g$  at time  $t$ . Here,  $G$  is defined as the time period when a unit first becomes treated, and  $C$  indicates the variable for whether the unit is in a never-treated group.

TABLE 1

*Graduate Record Examinations Test-Requirement Status Over Time*

Test Status	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Total
Required	99	96	95	101	101	97	96	94	95	95	83	70	1,122
Optional	0	4	5	5	5	10	10	13	14	17	29	39	151
Total	99	100	100	106	106	107	106	107	109	112	112	109	1,273

Callaway and Sant’Anna (2021) offer various ways to aggregate group-time average treatment effects across cohorts, overall units, and relative time-to-treatment assignment. For the estimation of the cohort-specific treatment effect, the following parameter is considered:

$$\theta_s(g) = \frac{1}{T-g+1} \sum_{t=2}^T 1\{g \leq t\} \text{ATT}(g,t), \quad (2)$$

which presents the average effect of treatment among units in the group  $g$ , across all post-treatment periods when there are  $T$  total time periods.<sup>2</sup> Built upon Equation (2), the overall average treatment effect can be estimated by averaging the  $\theta_s(g)$  together across groups, such that

$$\theta_s^o(g) := \sum_{g=2}^T \theta_s(g) P(G=g). \quad (3)$$

We further estimate the dynamic treatment effect to highlight treatment effect by relative time to the policy adoption.  $\text{ATT}(g,t)$  with respect to event-time  $e$  is

$$\theta_D(e) := \sum_{g=2}^T 1\{g+e \leq T\} \text{ATT}(g,g+e) P(G=g | G+e \leq T). \quad (4)$$

While the estimates from Equation (4) are useful for visually assessing whether the conditional parallel trend assumption holds, the pre-trend estimates produced by Callaway and Sant’Anna’s (2021) estimator need to be adjusted (Roth, 2024). By default, their specification calculates pretreatment period estimates relative to a “varying base period,” which is the immediately preceding period, rather than a “universal base period” (lead 1), which refers to the specific period immediately before treatment starts (Callaway, 2021). To address this, we also report transformed coefficients estimated relative to the baseline (lead 1) using the “eventbaseline” command in Stata 18

(Supplemental Appendix Tables 3 and 5 in the online version of the journal). For the event-study plots, however, we present those originally produced by Callaway and Sant’Anna’s (2021) estimator, as it is more effective in detecting anticipation effects (Callaway, 2021). In our main analyses, we focus on the parameters  $\theta_s^o(g)$  and  $\theta_D(e)$  as they are particularly useful for identifying the overall impact of policy adoption and the evolution of treatment effects over time.

To check the sensitivity of estimates, we conducted a robustness check using the estimator proposed by de Chaisemartin and D’Haultfoeuille (2024). Their approach is flexible enough that it is also applicable to fuzzy designs where treatment status is dynamic, although the actual treatment status of our sample is static. It calculates a weighted average of the DiD of units joining and leaving the treatment using not-yet-treated groups as the control group. Depending on the design and assumptions, the overall average treatment effect can be calculated in two ways. When assuming that past treatments do not affect current outcomes (i.e., no dynamic treatment effects), the average treatment effect corresponds to the average instantaneous effect (*event time*=0) and is equivalent to Callaway and Sant’Anna’s (CS) estimator without covariates using not-yet-treated units as controls (de Chaisemartin & D’Haultfoeuille, 2023). However, when dynamic treatment effects are present, it is more appropriate to calculate the average cumulative treatment effect over post-treatment periods (de Chaisemartin & D’Haultfoeuille, 2024). We report the estimates from the latter, as it is more reasonable to assume that changes in admissions policy generate delayed and accumulating effects on applicant behavior and programmatic decisions. The estimates of de Chaisemartin and D’Haultfoeuille (dCdH) models are reported in Supplemental Appendix Tables 2, 3, 4, and 6 (available in the online version of this article).

In addition, we included covariates in the analytic models using a regression adjustment approach. First, we used a proxy for unobserved confounders that were likely to be highly correlated with the program-specific trend. By including the proxy variable, we can reduce the bias in coefficients that could be confounded by unobserved characteristics (Bollinger & Minier, 2015). Specifically, we used information about the number of PhD degrees granted nationally in each 6-digit Classification of Instructional Programs (CIP) code from the Integrated Postsecondary Education Data System as the proxy. CIP code is a taxonomy of academic programs developed by the U.S. Department of Education. This variable controls for the cross-field national trends in demand for getting a PhD in each program.

Next, we used enrollment in an institutional fellowship program awarded to the admitted students who come from diverse cultural, geographic, and family backgrounds to control for the program's commitment and capacity to recruit students from disadvantaged backgrounds that are underrepresented in the field of study. Students do not apply directly for the fellowship but are nominated by their admitting graduate programs. Recipients of this fellowship receive full funding for 5 years, with costs shared between the graduate school and the program. Non-recipients are also guaranteed full funding for 5 years at most programs, and for 4 years at a few programs, subsidized entirely from their programs. The cost-sharing system incentivizes graduate programs to recruit students from historically underrepresented backgrounds. We generated two variables to represent the share of students nominated for and receiving the fellowship in each program per year. While the inclusion of the fellowship variables helps account for program-level diversity efforts that may otherwise confound the estimates, it could be endogenous if it is correlated with the treatment. To address this concern, we provide the results of additional checks excluding this control and find that our main results remain robust (see Supplemental Appendix Tables 2 and 4 in the online version of the journal).

Lastly, we conducted an additional robustness check by excluding the observations from the 2019–2020 admissions cycle. While the

application and admission decisions for the 2019–2020 cycle occurred prior to the onset of the COVID pandemic and were not affected by its disruptions, the pandemic might have had a substantial impact on yield. To eliminate the potential COVID effects on the 2019–2020 admissions cycle results, the adjusted sample was used, and the estimates were not substantively different from the original estimates in terms of significance, magnitude, and direction (see Supplemental Appendix Tables 2, 3, and 4 in the online version of the journal).

*Limitations.* Several limitations of this study should be noted. First, our findings are not applicable for making inferences about discipline-specific features. Our sample is derived from a research-intensive university that offers one of the broadest ranges of PhD programs in the United States. Although we leveraged this variability to explore general mechanisms in PhD admissions while controlling for disciplinary differences by including school-level fixed effects, our results do not speak to specific disciplinary logics at the program or department level. Despite improved internal validity with the inclusion of fixed effects, we acknowledge variation across disciplines in the timing of policy adoption. GRE-optional policies were adopted earlier by programs in Humanities and Biological Sciences, whereas adoption in Engineering and Social Sciences occurred later, largely clustered around 2018 and beyond (see Supplemental Appendix Table 5 in the online version of the journal). The determinants of policy adoption may have varied across disciplines over time. Thus, extrapolating our study findings to specific disciplinary contexts should be done with caution.

Next, our analytical sample is limited to admissions records from 2009 to 2020, which predates the COVID-19 pandemic. This focused time frame allows us to make robust inferences about the policy impact while isolating the confounding effects of the pandemic on admissions. However, given the significant expansion of test-optional policies triggered by the pandemic, the motivations for adopting these policies in our sample likely differ from those of programs that became test-optional during the pandemic. Pre-pandemic adoption was mainly voluntary, while post-pandemic adoption was more likely to be

TABLE 2

*Overall ATT of Graduate Record Examinations -Optional Policies on Application Growth, Acceptance Rate, and Yield Rate*

Outcomes	Application growth rate	Acceptance rate	Yield rate
Overall	0.175* (0.074)	-0.067* (0.033)	0.029 (0.040)
<i>N</i>	1,273	1,242	1,227

*Note.* Standard errors are clustered at the program level and reported in parentheses. All specifications include covariates and fixed effects. Coefficients were estimated using Callaway and Sant’Anna’s (2021) approach and the “xtdidregress” commands in Stata.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

involuntary (Bastedo et al., 2025). Further, treatment adoption is concentrated in the late 2010s, and the aggregated estimates are likely to be heavily weighted by adoption from this period. Thus, the contexts in which the GRE-optional policies were adopted should be considered when interpreting our results.

Lastly, our sample is highly representative of programs that are often contained within elite institutions. The curriculum, education goals, student and faculty profiles of these programs may look different from those of corresponding programs at other institutions. Thus, our findings might not be generalizable to institutions with different profiles of disciplines or not focused on intensive research.

## Findings

### *Treatment Effects on Enrollment Outcomes*

To answer our first research question, we employed DiD methods to examine the impact of GRE-optional policies on application counts, selectivity as measured by acceptance rates, and yield rates. Table 2 presents the estimates for the programs included in our analysis. Given the staggered policy adoption in our study setup, we discuss results that are consistent in statistical significance, magnitude, and direction of estimates across the heterogeneity-robust DiD estimators suggested by CS and dCdH. All estimates from the traditional TWFE, CS, and dCdH models are presented in Supplemental Appendixes (available in the online version of this article).

The findings indicate that the number of applications grew 17.5% after GRE requirements were eliminated. Given relatively stable admitted

cohort sizes, this increase in applications contributes to selectivity by lowering acceptance rates by 6.7 percentage points (%pts, hereafter). There was a slight increase in the yield rate, albeit not statistically significant.

To investigate whether the treatment effects of GRE-optional policies on applications and acceptance rates remain constant over time, we estimated their dynamic effects (Supplemental Appendix Table 3 in the online version of the journal). This event-study approach also presents the findings of a visual and statistical check of whether the parallel trends assumption is satisfied. This assumption needs to be met to make a valid causal inference about treatment effects. We plotted the pointwise estimates of treatment effects across the length of exposure to the GRE-optional policies in Figure 1. The findings indicate that the parallel trend assumption is violated for the application growth outcome. There is an upward trend that begins a year prior to the elimination of GRE requirements, and its point estimate is statistically significant. This violation is apparent in the adjusted estimates produced by Roth’s estimator (see Supplemental Appendix Table 3 in the online version of the journal). Moreover, the significant effects are mostly instantaneous and no long-term effects are evident. There is a positive effect of GRE-optional policies on application growth immediately in the same year the policies are adopted, but the effect vanishes soon after the adoption year.

In contrast, the parallel trends assumption is satisfied for the acceptance rate. Although the alternative estimator (dCdH) has one significant estimate at lead 3, all other pretreatment leads

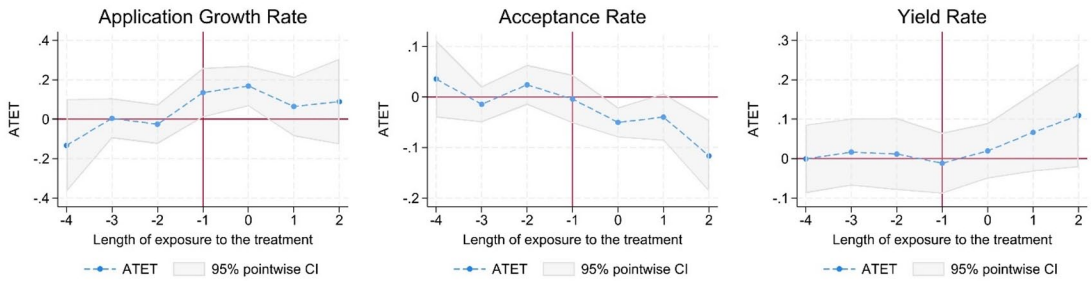


FIGURE 1. *Dynamic treatment effect estimates of Graduate Record Examinations-optional policies on application growth, selectivity, and yield using Callaway and Sant'Anna's (2021) estimator.*

TABLE 3

*Overall Average Treatment Effects of Graduate Record Examinations-Optional Policies on Structural Diversity and UGPA*

Panel	Female	URM	International	Pell	First-gen	UGPA
Application	0.089*** (0.028)	0.015 (0.027)	-0.009 (0.023)	0.033 (0.048)	0.004 (0.028)	0.033 (0.038)
Admission	0.095 (0.059)	-0.075 (0.047)	0.013 (0.041)	-0.040 (0.051)	-0.032 (0.039)	0.174 (0.095)
Enrollment	0.096 (0.109)	0.022 (0.039)	-0.022 (0.047)	-0.048 (0.057)	-0.017 (0.049)	0.208 (0.129)

*Note.* Standard errors are clustered at the program level and reported in the parentheses. All specifications include covariates and fixed effects. Coefficients were estimated using Callaway and Sant'Anna's (2021) approach and the "xthdidiress" commands in Stata. URM=underrepresented racially/ethnically minoritized; UGPA=undergraduate GPA.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

remain not significant. The joint null hypothesis of the pre-trend coefficients is not statistically significant, thereby supporting the satisfaction of the assumption (see Supplemental Appendix Table 3 in the online version of the journal). We thus interpret this isolated significance as random noise rather than a systematic pre-trend. The findings show that the acceptance rate decreases as soon as policies are adopted and further decreases 2 years thereafter. In terms of yield rate, the parallel trend assumption is satisfied, but no significant dynamic effects are found.

#### *Treatment Effects on Structural Diversity*

We estimated the treatment effect on the composition of the student body to answer research question 2. The average treatment effects are provided in Table 3. Our findings indicate that GRE-optional policy adoption does not significantly change applicant or admitted student demographics except the share

of female students who apply, which did not result in an increased share of female enrollees. The share of female students increased by 8.9%pts among the applicant pool, but the increase of females in the admitted and enrollee pools were not statistically significant.

We also estimated the dynamic treatment effects of policy adoption on structural diversity at each stage of the admissions process: application, admission, and enrollment (Figures 2, 3 and 4; Supplemental Appendix Table 6 in the online version of the journal). The results demonstrate an immediate increase in the proportion of female applicants following the transition to GRE-optional policies in PhD admissions. The instantaneous effect lasted for a year and was not statistically significant in the third year after policy adoption. Unlike the applicant pool, a significant change in the proportion of women in the admitted and enrollee pools was not observed.

We found additional evidence of dynamic GRE-optional treatment effects on structural

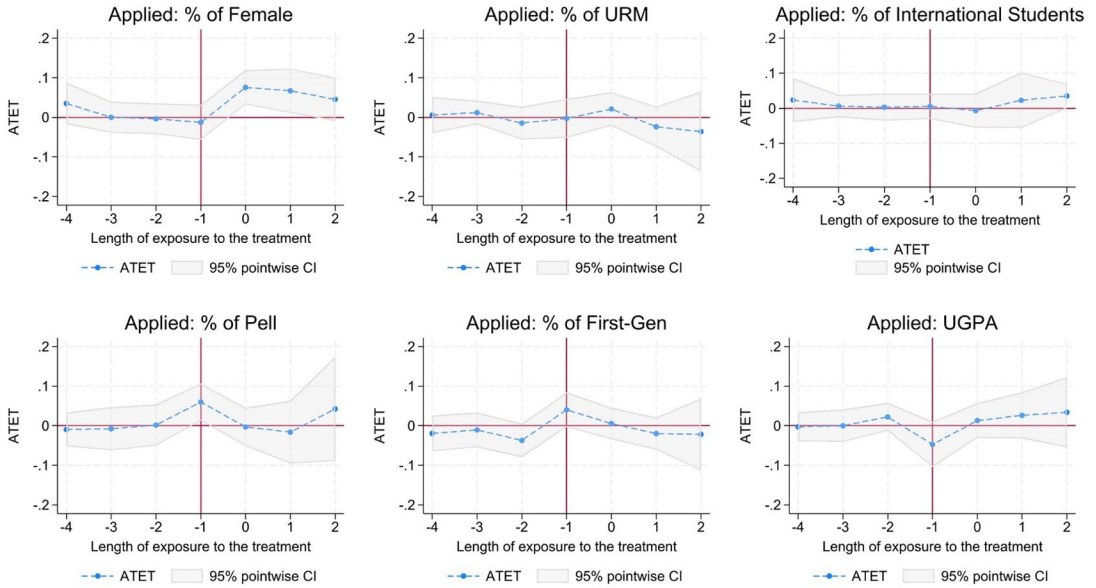


FIGURE 2. *Dynamic treatment effect estimates on the structural diversity of applicants.*

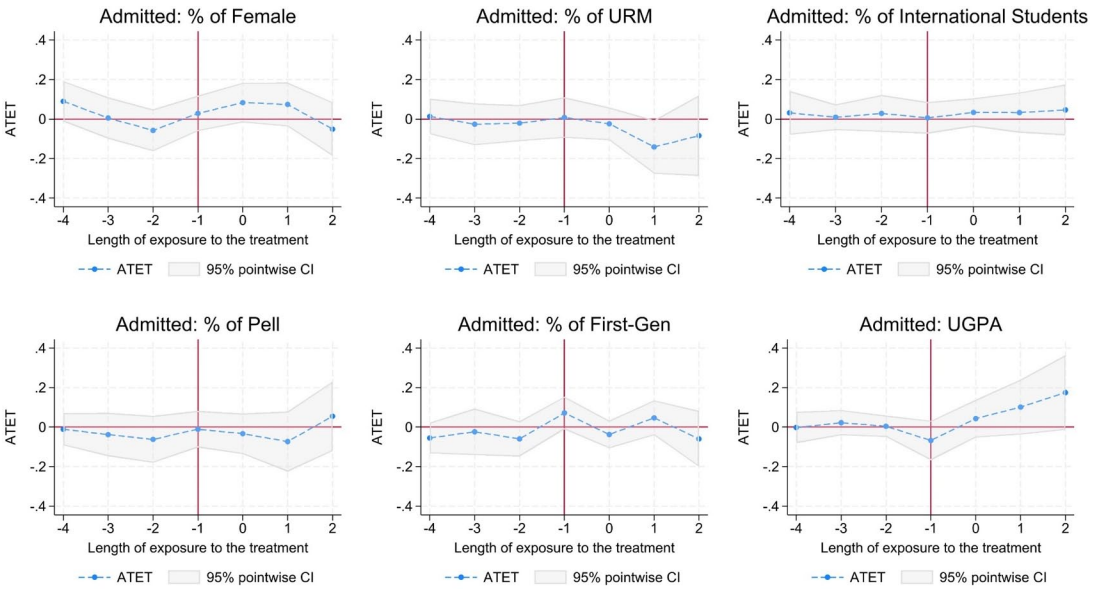


FIGURE 3. *Dynamic treatment effect estimates on the structural diversity of admitted applicants.*

diversity, particularly at the admission stage. Specifically, the share of racially minoritized students among admitted applicants decreased in the year following the programs' shift away from requiring GRE scores, as shown in Figure 3. This negative effect becomes no longer statistically significant 2 years after adoption. Enrollment

(yield) of racially minoritized students was unaffected by GRE-optional policy adoption.

Regarding other outcomes, no significant changes were found after the adoption of GRE-optional policies. The parallel trend assumption was compromised among the outcomes, notably the proportions of Pell Grant recipients and

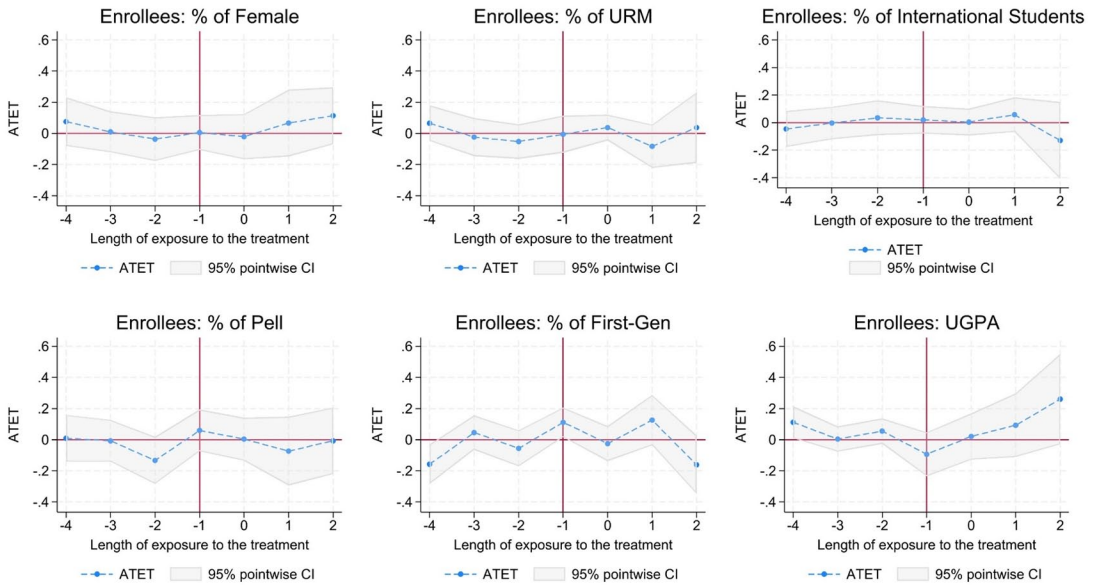


FIGURE 4. *Dynamic treatment effect estimates on the structural diversity of enrollees.*

first-generation college students, due to the significant pre-trend estimates of lead 1, a year before the policy adoption (see Supplemental Appendix Table 6 in the online version of the journal). In terms of the average UGPA scores, we found an increase in the GPA of admitted and enrolled pools following the change in GRE policies. While the post-trend estimates for this increase were not statistically significant, the magnitudes were substantial and increasing over time, accompanied by wide confidence intervals. The large confidence intervals indicate that programs have a large variation in how they weigh undergraduate GPA in making admissions decisions after implementing GRE-optional policies.

### Discussion and Implications

The results of this study indicate that removing GRE score requirements enhances institutional selectivity, as shown by drops in acceptance rates. Although this effect was noticeable in the first year following the policy change, it did not persist over time. This finding aligns with existing research on undergraduate college admissions (Belasco et al., 2015; Rubin & González Canché, 2019), which also observed that such effects do not last (Saboe & Terrizzi, 2019).

Regarding the violation of the parallel trend assumption on application growth, we interpret it as an indicator of an institutional climate in which test-optional policies may have reinforced existing momentum. Although the parallel trends assumption did not hold for this outcome, the significant effect observed in the year of policy adoption, coupled with the increase in pre-trend, suggests a potential anticipation effect. This trend implies that prospective applicants or programs may have responded to informal signals, discussions, or early announcements of upcoming policy changes before the official implementation of new policies. Such pretreatment shifts also suggest that enhanced selectivity may not be solely attributable to the formal adoption of GRE-optional policies, but rather partially due to changes in disciplines, institutions, or the prospective student pool already underway.

Regarding structural diversity, the primary outcome of interest in this study, GRE-optional policies did not significantly change student body composition. Our findings show an absence of substantial shifts in demographics among enrollees. Notably, racially minoritized students were less likely to be admitted in the year after policy adoption, but this did not negatively impact the enrolled pool. Although the application pool diversified slightly by gender, evidenced by an increase

in female applicants, the final composition of the student body remained unchanged after admission decisions were made. This outcome aligns with previous research indicating that adopting test-optional policies alone is insufficient for improving structural diversity (Conlin & Dickert-Conlin, 2017; Rosinger et al., 2022; Saboe & Terrizzi, 2019), but contrasts with recent research finding more positive effects at private undergraduate institutions (Bennett, 2022).

While many of the estimated treatment effects on diversity-related outcomes are not statistically significant, they may still have practical significance. In our findings, for instance, GRE-optional policies were associated with an increase in UGPAs, particularly among admitted applicant and enrollee pools. The high magnitude of the coefficient, nearly a fifth of a grade point on a 4.0 scale, suggests that programs without the test requirement may have admitted applicants with higher undergraduate GPAs. While not statistically significant, this result indicates a greater weight programs may have placed on the conventional academic measures to assess academic preparedness in the absence of GRE scores. In addition, the relatively large increase in female enrollment, though not statistically significant, suggests possible heterogeneous effects by discipline that may have been masked in overall estimates. Given the large standard errors and the small sample size within fields that limit drawing robust inference (Supplemental Appendix Tables 7 and 8 in the online version of the journal), further research is needed to evaluate the treatment impact by discipline.

Another factor to consider is the effect of other institutions joining the test-optional movement. Programs' decisions to remove standardized test scores from doctoral admissions did not occur in isolation, but as part of a broader context of a nationwide shift (GRExit) that emerged in the 2010s (King et al., 2020). For instance, nearly half of molecular biology PhD programs had stopped requiring GRE scores by 2018, and other disciplines such as History, Public Health, and Life Sciences at several prominent institutions (e.g., Yale, Boston University) also removed the test requirement in doctoral admissions (Association of Schools and Programs of Public Health, 2023; King et al., 2020; Langin, 2019, 2022; Sullivan et al., 2022; Yale University Department of

History, 2019). This historical context suggests policy diffusion that an institution's admissions policies are influenced by peer institutions that are agents of reform in doctoral admissions (Furuta, 2017). Such interdependence among institutions is particularly salient in the higher education system, where an institution's change in its admissions policy informs other institutions about the details and outcomes of it, thereby generating the feedback loops of learning for organizational change (Kim, 2024). Thus, applicant behavior and each program's decision-making are likely shaped not only by internal policy deliberations but also by external factors. While this dynamic enhances the external validity of this study, the treatment effect needs to be considered within a broader context of contemporaneous shifts that may have influenced changes in applicant and/or institutional behaviors.

One of our study's major contributions is its examination of test-optional policy impacts across stages in the admissions process, allowing for a better understanding of the potential mechanisms driving test-optional policies at the graduate program level. Applying our conceptual framework to the study results, we see evidence that the impact of GRE-optional policies primarily operates through a short-lived "publicity" mechanism at the application stage. Contrary to a major rationale for implementing test-optional and test-free policies, removing GRE requirements does not appear to disproportionately lower application barriers for students marginalized by race or social class. Even among women, whose application rates after policy implementation increased faster than men, the increase was short lived. Looking at yield rates, we see no evidence that GRE-optional policies operate through a "signaling" mechanism by prompting admitted marginalized students to view their fit with the institution more positively.

Looking at the admitted applicant pool, our findings suggest that risk-aversion among faculty remained potent and that simply adopting test-optional policies, on their own, is unlikely to generate significant impacts on the diversity of PhD students. Effective shifts in faculty evaluations toward a more holistic and contextual review, in addition to substantial recruitment efforts, have the potential to diversify admitted and enrolled students (Posselt et al., 2023).

Without these corresponding efforts, however, simply changing testing policies is most likely to reproduce existing conditions. Variation in program implementation also likely plays a significant role. Bastedo et al. (2025) found, for example, that the way undergraduate test-optional policies were implemented varied greatly across institutions during the pandemic. For instance, many admissions offices continued to rely on traditional evaluation methods based on numerical metrics of academic potential rather than engaging in more holistic forms of review. Some institutions created dual evaluation systems for applicants with and without test scores or even developed quantitative indexes based on GPA and AP scores. Doctoral admissions committees often identify candidates whose backgrounds, experiences, and interests align with their programs' existing and emerging priorities (Posselt, 2015). This homophily leads to a consistently similar student body in terms of sociodemographic and academic prestige (Posselt, 2016). Without comprehensive multilevel changes, including policy alterations, cultural shifts, and new cognitive routines, the transformative potential of test-optional admissions to disrupt structural inequities remains compromised (Baker & Rosinger, 2020; Bastedo et al., 2025; Posselt et al., 2023).

This study has important implications for doctoral admissions practices. Faculty members tend to be risk-averse in doctoral admissions due to uncertainties about applicants' academic potential and the high-stakes nature of the admissions process. Thus, programs introducing GRE-optional policies should incorporate multilevel interventions to make faculty more comfortable in an admissions review process that promotes inclusive excellence. Implementing successful GRE-optional policies requires adopting more holistic and individualized approaches to the admissions process (Posselt et al., 2023). This is challenging as it involves changing established evaluation traditions and handling the increased complexity and time demands of the admissions process (Kent & McCarthy, 2016). To ease this transition, programs can support faculty by offering targeted training, clear guidelines, and rubrics for reviewing application materials, which helps faculty members avoid prematurely rejecting strong candidates

(Kent & McCarthy, 2016; Posselt et al., 2023). For example, workshops focused on improving inter-rater reliability for holistic review can enhance both the efficiency and integrity of decision-making processes (Addams et al., 2010; American Association of Colleges of Nursing, 2020). In addition, faculty should receive comprehensive information to assess applications within the contexts of prior educational challenges, experiences, and opportunities. With more comprehensive information about applicants' achievements relative to their educational opportunities, evaluators can better reflect on their norms, values, and heuristics used during application evaluations (Bastedo et al., 2022; Pieper & Krsmanovic, 2023). Training opportunities can be provided by the scholarly community to help faculty within their disciplines establish shared evaluative practices that mitigate the influence of implicit bias in admissions decisions (Association of Schools and Programs of Public Health, 2023).

Program leaders need to provide systematic support to faculty members, alleviating concerns about financial, administrative, and academic implications. Posselt et al. (2023) emphasize that implementing holistic review requires navigating cultural changes and addressing the perceived legitimacy of specific criteria. Leaders must shape priorities, manage colleagues' concerns or resistance, and "empower faculty to discuss racism, racialized organizational cultures, and implications for policy and practice" (Posselt et al., 2023, p. 19). By embedding the core concepts of holistic review into organizational practices, faculty would feel less risk in conducting admissions practices that enhance structural diversity. Furthermore, concurrent investments in recruitment and financial aid can reduce the risk perceived by faculty in student selection. Although this varies by institution or discipline, faculty frequently support PhD students with their funding, leading them to view admissions as an economic investment. Given the time and financial costs tied to student recruitment, faculty tend to select applicants who appear to be the safest choices in terms of having backgrounds and credentials that mirror the faculty's own. Hence, institutional support should include funding programs to recruit students from diverse backgrounds and improving program climate and policies to support adviser-advisee relationships.

Targeted recruitment efforts at the institutional level could alleviate such risk from faculty in admitting students from underrepresented backgrounds by identifying students with academic potential who might otherwise be unrecognized. The Fisk-Vanderbilt Master's-to-PhD Bridge Program, for example, is an initiative in which a historically Black university partners with a leading research institution to provide a structured pathway to STEM doctoral programs. Participants are provided with rigorous training in coursework and intensive research experiences, which enable them to demonstrate their potential for success in doctoral study (Stassun et al., 2011). Given the program's high quality and its role in reducing information gaps, faculty can make decisions about these applicants with less uncertainty. Furthermore, a streamlined pipeline from outreach to admission needs to be established. Despite the limited role administrators can actually serve on graduate admission committees, they can influence admissions by advising committees about promising applicants from underrepresented groups (Griffin & Muñiz, 2015). Such broader institutional commitment can help reshape structural diversity by strengthening the impact of recruitment efforts on the admissions stage.

Debates over the value of standardized tests in admissions have often been rooted in anecdotal narratives and political discourse. This study contributes empirical evidence about the impact of GRE-optional policies on applications, yields, and structural diversity using quasi-experimental methods that reflect recent advances in this discipline. This research extends the scholarship of holistic admissions to the graduate level, elucidating the often opaque process of PhD admissions and revealing the potential of admissions committees to act as gatekeepers if commitments to inclusive excellence are not institutionalized. We offer insights into the mechanisms of PhD admissions and identify where barriers to access for marginalized students persist. This research encourages future studies to explore how student and institutional behaviors may yield different enrollment outcomes under test-free policies (Lo et al., 2023) and how strategic disclosure of test scores by doctoral applicants may impact their

probability of admission (McManus et al., 2023). Additional research on GRE-optional policies and structural diversity would also benefit from an extended sample drawn from post-pandemic periods (Rosinger et al., 2024).

### Acknowledgments

The authors gratefully acknowledge the Rackham Graduate School for its generous support of this study. We also thank Kevin Stange, Brian Jacob, Akila Weerapana, Aashish Mehta, Abbi Crowder, Paula Clasing-Manquian, participants at CIERS, the AEFPP 2024 Conference, and the ASSA 2025 Conference, as well as the anonymous reviewers for their valuable feedback. Any errors, omissions, findings, conclusions, or recommendations expressed in this article are solely those of the authors.

### Declaration of Conflicting Interests


The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Some of the authors (Kim, Gonzalez, Flaster, and Bai) are or were previously funded by the Rackham Graduate School. The authors received no additional financial benefit directly related to the authorship or publication of this article from the school.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Open access funding for this article was provided by the University of Houston (R0511911).

### ORCID iDs

Heeyun Kim  <https://orcid.org/0000-0002-8742-6287>

Michael Bastedo  <https://orcid.org/0000-0003-3358-2564>

Allyson Flaster  <https://orcid.org/0000-0002-8529-9130>

### Notes

1. Test-free admissions do not allow students to submit standardized test scores for consideration, whereas test-optional admissions grant students the choice to submit scores or not (Park et al., 2024).

2. The term involving the indicator function,  $\mathbf{1}\{\cdot\}$ , limits the estimation to the identified units. Otherwise, it returns 0.

## References

- Addams, A. N., Bletzinger, R. B., Sondheimer, H. M., White, S. E., & Johnson, L. M. (2010). *Roadmap to diversity: Integrating holistic review practices into medical school admission processes*. Association of American Medical Colleges.
- Agan, A., & Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics*, *133*(1), 191–235.
- American Association of Colleges of Nursing. (2020, February 27). *Establishing inter-rater reliability for holistic admissions review* [Webinar]. [https://www.aacnursing.org/conferences-webinars/info/sessionaltcd/WF20\\_02\\_27](https://www.aacnursing.org/conferences-webinars/info/sessionaltcd/WF20_02_27)
- Association of Schools and Programs of Public Health. (2023, October 26). *ASPPH presents webinar: Implicit bias in application review training*. <https://aspqh.org/aspqh-presents-webinar-implicit-bias-in-application-review-training/>
- Baker, D., & Rosinger, K. (2020). Test optional offers benefits but it's not enough. *Education Next*, *20*(3), 67–70.
- Bastedo, M. N., & Bowman, N. A. (2017). Improving admission of low-SES students at selective colleges: Results from an experimental simulation. *Educational Researcher*, *46*(2), 67–77.
- Bastedo, M. N., Glasener, K. M., Deane, K. C., & Bowman, N. A. (2022). Contextualizing the SAT: Experimental evidence on college admission recommendations for low-SES applicants. *Educational Policy*, *36*(2), 282–311.
- Bastedo, M. N., & Jaquette, O. (2011). Running in place: Low-income students and the dynamics of higher education stratification. *Educational Evaluation and Policy Analysis*, *33*(3), 318–339.
- Bastedo, M. N., Kapp, R., Bai, Y., & Carroll, S. (2025). Dodging the competency trap? How admissions and financial aid offices adapted to test-optional policies after COVID-19. *The Journal of Higher Education*, 1–25. <https://doi.org/10.1080/00221546.2024.2446014>
- 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.
- Bennett, C. T. (2022). Untested admissions: Examining changes in application behaviors and student demographics under test-optional policies. *American Educational Research Journal*, *59*(1), 180–216.
- Benshoff, J. M., Cashwell, C. S., & Rowell, P. C. (2015). Graduate students on campus: Needs and implications for college counselors. *Journal of College Counseling*, *18*(1), 82–94.
- Bersola, S. H., Stolzenberg, E. B., Love, J., & Fosnacht, K. (2014). Understanding admitted doctoral students' institutional choices: Student experiences versus faculty and staff perceptions. *American Journal of Education*, *120*(4), 515–543.
- Bleske-Rechek, A., & Browne, K. (2014). Trends in GRE scores and graduate enrollments by gender and ethnicity. *Intelligence*, *46*, 25–34.
- Bollinger, C. R., & Minier, J. (2015). On the robustness of coefficient estimates to the inclusion of proxy variables. *Journal of Econometric Methods*, *4*(1), 101–122.
- Cahn, P. S. (2015). Do health professions graduate programs increase diversity by not requiring the graduate record examination for admission? *Journal of Allied Health*, *44*(1), 51–56.
- Callaway, B. (2021, November). *Universal vs. varying base period in event studies*. <https://bcallaway11.github.io/posts/event-study-universal-v-varying-base-period>
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.
- Cheslock, J. J., & Kroc, R. (2012). Managing college enrollments. *The handbook for institutional researchers*, 221–236.
- Conlin, M., & Dickert-Conlin, S. (2017). Inference by college admission departments. *Journal of Economic Behavior & Organization*, *141*, 14–28.
- Dang, K. V., Rerolle, F., Ackley, S. F., Irish, A. M., Mehta, K. M., Bailey, I., Fair, E., Miller, C., Bibbins-Domingo, K., Wong-Moy, E., Glymour, M. M., & Morris, M. D. (2021). A randomized study to assess the effect of including the graduate record examinations results on reviewer scores for underrepresented minorities. *American Journal of Epidemiology*, *190*(9), 1744–1750.
- de Chaisemartin, C., & D'Haultfœuille, X., (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.
- de Chaisemartin, C., & D'Haultfœuille, X. (2023). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal*, *26*(3), C1–C30.
- de Chaisemartin, C., & D'Haultfœuille, X. (2024). *Difference-in-differences estimators of intertemporal treatment effects*. National Bureau of Economic Research. <https://doi.org/10.3386/w29873>
- DesJardins, S. L., Ahlburg, D. A., & McCall, B. P. (2006). An integrated model of application, admission, enrollment, and financial aid. *The Journal of Higher Education*, *77*(3), 381–429.
- Elias, J. (2022, November 16). Test-optional policies now dominate higher ed. *The Chronicle of Higher*

- Education. <https://www-chronicle-com/article/test-optional-policies-now-dominate-higher-ed>.
- Epstein, J. P. (2009). Behind the SAT-optional movement: Context and controversy. *Journal of College Admission, 204*, 8–19.
- Feldon, D. F., Litson, K., Cahoon, B., Feng, Z., Walker, A., & Tofel-Grehl, C. (2024). The predictive validity of the GRE across graduate outcomes: A meta-analysis of trends over time. *The Journal of Higher Education, 95*(1), 120–148.
- Friedman, J. N., Sacerdote, N. B., & Tine, N. M. (2024). *Standardized test scores and academic performance at Ivy-plus colleges*. Opportunity Insights. [https://opportunityinsights.org/wp-content/uploads/2024/01/SAT\\_ACT\\_on\\_Grades.pdf](https://opportunityinsights.org/wp-content/uploads/2024/01/SAT_ACT_on_Grades.pdf)
- Furuta, J. (2017). Rationalization and student/school personhood in US college admissions: The rise of test-optional policies, 1987 to 2015. *Sociology of Education, 90*(3), 236–254.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics, 225*(2), 254–277.
- Griffin, K. A., & Muñoz, M. (2015). Rethinking the structure of student recruitment and efforts to increase racial and ethnic diversity in doctoral education. *International Journal of Doctoral Studies, 10*, 199.
- Hernández-Colón, I. R., Caño, A., Wurm, L. H., Sanders, G., & Nava, J. (2021). Instructional set moderates the effect of GRE on faculty appraisals of applicant competence: A vignette study with implications for holistic review. *Frontiers in Psychology, 12*, Article 749621.
- Hooker, E. A., Mallow, P., Downes, C., & Baidwan, S. (2022). Abandoning the GRE and GMAT improved diversity in a graduate program for health administration. *Journal of Health Administration Education, 38*(4), 895–912.
- Kallio, R. E. (1995). Factors influencing the college choice decisions of graduate students. *Research in Higher Education, 36*, 109–124.
- Kell, H., Roohr, K. C., & Cline, F. A. (2021). *What is the relationship between demographic background and test score submission to GRE-optional graduate programs? It's complicated*. Education Testing Service.
- Kent, J. D., & McCarthy, M. T. (2016). *Holistic review in graduate admissions*. Council of Graduate Schools.
- Kim, H. (2024). Bridging principal-agent and mechanism design theories: An integrated conceptual framework for policy evaluation. *Asia Pacific Education Review, 25*(2), 329–342.
- King, M. R., Jennings, G. K., Chalkley, R. G., & Sealy, L. J. (2020). Questioning the value of the graduate record examinations (GRE) in PhD admissions in biomedical engineering. *Annals of Biomedical Engineering, 48*, 2155–2157.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127*(1), 162–181.
- Langin, K. (2019). *A wave of graduate programs drops the GRE application requirement*. Science. <https://www.science.org/content/article/wave-graduate-programs-drop-gre-application-requirement>
- Langin, K. (2022). Pandemic accelerates “GRExit.” *Science, 378*(6623), 936–937.
- Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision-making analysis. *Organizational Behavior and Human Decision Processes, 69*(2), 149–163.
- Lo, P., Park, J. J., Wong, N., Zheng, J., Poon, O. A., & Rosinger, K. (2023). *Test-free admissions at selective colleges: Insights from admissions professionals*. College Admissions Futures Co-Laborative. <https://www.caicolab.org>
- March, J. G. (1994). *A primer on decision making: How decisions happen*. Simon and Schuster.
- March, J. G., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science, 33*(11), 1404–1418.
- McManus, B., Howell, J., & Hurwitz, M. (2023). *Strategic disclosure of test scores: Evidence from U.S. college admissions* (EdWorkingPaper No. 23-843). Annenberg Institute for School Reform at Brown University.
- Millar, J. A. (2020). The GRE in public health admissions: Barriers, waivers, and moving forward. *Frontiers in Public Health, 8*, Article 609599.
- Miller, C. W., Zwickl, B. M., Posselt, J. R., Silvestrini, R. T., & Hodapp, T. (2019). Typical physics Ph. D. admissions criteria limit access to underrepresented groups but fail to predict doctoral completion. *Science Advances, 5*(1), Article eaat7550.
- National Center for Science and Engineering Statistics. (2024). *Academic institution profiles – Rankings by earned doctorates*. Retrieved November 19, 2025, from <https://ncesdata.nsf.gov/profiles/site?method=rankingBySource&ds=drf>
- Nicklin, J. M., & Roch, S. G. (2008). Biases influencing recommendation letter contents: Physical attractiveness and gender. *Journal of Applied Social Psychology, 38*(12), 3053–3074
- Office of the Vice President for Research, University of Michigan. (2024, October 24). *FY24 research annual report*. <https://research.umich.edu/news-and-issues/research-annual-reports/fy24-research-annual-report/>

- Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics*, 33(2), 493–520.
- Paris, J. H., Torsney, B., Fiorot, S., & Beckowski, C. P. (2022). The impact of optional: Investigating the effects of test-optional admissions policies. *Journal of College Access*, 7(2), 7–29.
- Park, J. J., Wong, N., Lo, P., Zheng, J., Poon, O., & Rosinger, K. (2024). Test-free admissions at selective institutions: Perspectives from admissions professionals. *Journal of College Student Development*, 65(4), 438–442.
- Pieper, B., & Krsmanovic, M. (2023). Faculty perceptions on (implicit) bias during the graduate admission review process. *Studies in Graduate and Postdoctoral Education*, 14(2), 117–133.
- Poock, M. C., & Love, P. G. (2001). Factors influencing the program choice of doctoral students in higher education administration. *NASPA Journal*, 38(2), 203–223.
- Posselt, J. R., Jaquette, O., Bielby, R., & Bastedo, M. N. (2012). Access without equity: Longitudinal analyses of institutional stratification by race and ethnicity, 1972–2004. *American Educational Research Journal*, 49(6), 1074–1111.
- Posselt, J. R. (2014). Toward inclusive excellence in graduate education: Constructing merit and diversity in PhD admissions. *American Journal of Education*, 120(4), 481–514.
- Posselt, J. R. (2015). Disciplinary logics in doctoral admissions: Understanding patterns of faculty evaluation. *The Journal of Higher Education*, 86(6), 807–833.
- Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gatekeeping*. Harvard University Press.
- Posselt, J. R., Southern, D., Hernandez, T., Desir, S., Alleyne, F., & Miller, C. W. (2023). Redefining merit through new routines: Holistic admissions policy implementation in graduate education. *Educational Evaluation and Policy Analysis*, 47(1), 159–184. <https://doi.org/10.3102/01623737231201612>.
- Rosinger, K. O., Baker, D. J., Sturm, J., Yu, W., Park, J. J., Poon, O., & Breen, S. (2024). *Exploring the relationship between test-optional admissions and selectivity and enrollment outcomes during the pandemic* (EdWorkingPaper No. 24-982). Annenberg Institute for School Reform at Brown University.
- Rosinger, K. O., Ford, K. S., Posselt, J., & Choi, J. (2022). Exploring the impact of GRE-accepting admissions on law school diversity and selectivity. *The Review of Higher Education*, 46(1), 109–144.
- Rosinger, K. O., Sarita Ford, K., & Choi, J. (2021). The role of selective college admissions criteria in interrupting or reproducing racial and economic inequities. *The Journal of Higher Education*, 92(1), 31–55.
- Roth, J. (2024). *Interpreting event-studies from recent difference-in-differences methods* (arXiv preprint arXiv:2401.12309).
- Rubin, P. G., & González Canché, M. S. (2019). Test-flexible admissions policies and student enrollment demographics: Examining a public research university. *The Review of Higher Education*, 42(4), 1337–1371.
- Saboe, M., & Terrizzi, S. (2019). SAT optional policies: Do they influence graduate quality, selectivity or diversity?. *Economics Letters*, 174, 13–17.
- Saul, S. (2023, May 31). *The GRE test is cut in half: Two hours and done*. The New York Times. <https://www.nytimes.com/2023/05/31/us/gre-test-shorter-graduate-school.html>.
- Schultz, L., & Backstrom, B. (2021). *Test-optional admissions policies: Evidence from implementations pre-and post-COVID-19*. Policy Brief. Nelson A. Rockefeller Institute of Government.
- Stassun, K. G., Sturm, S., Holley-Bockelmann, K., Burger, A., Ernst, D. J., & Webb, D. (2011). The Fisk-Vanderbilt Master’s-to-Ph. D. Bridge Program: Recognizing, enlisting, and cultivating unrealized or unrecognized potential in underrepresented minority students. *American Journal of Physics*, 79(4), 374–379.
- Sullivan, L. M., Velez, A. A., Longe, N., Larese, A. M., & Galea, S. (2022). Removing the graduate record examination as an admissions requirement does not impact student success. *Public Health Reviews*, 43, Article 1605023.
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124–1131.
- Wall Bortz, W. E., Knight, D. B., Lyles, C. H., Kinoshita, T., Choe, N. H., Denton, M., & Borrego, M. (2020). A competitive system: Graduate student recruitment in STEM and why money may not be the answer. *The Journal of Higher Education*, 91(6), 927–952.
- Woo, S. E., LeBreton, J. M., Keith, M. G., & Tay, L. (2023). Bias, fairness, and validity in graduate-school admissions: A psychometric perspective. *Perspectives on Psychological Science*, 18(1), 3–31.

- Xu, Y. J. (2014). Advance to and persistence in graduate school: Identifying the influential factors and major-based differences. *Journal of College Student Retention: Research, Theory & Practice*, 16(3), 391–417.
- Yale University Department of History. (2019, August 2). *A note on removing the GRE from graduate admissions*. Yale University. <https://history.yale.edu/news/note-removing-gre-graduate-admissions>
- Zhang, L. (2005). Advance to graduate education: The effect of college quality and undergraduate majors. *The Review of Higher Education*, 28(3), 313–338.

### Authors

HEEYUN KIM, PhD, is an assistant professor at the University of Houston, and part of this study was conducted while she was a research fellow at the University of Michigan. Her research interests include admissions policies and graduate education, with a focus on leveraging institutional, state, and federal data to evaluate policy impacts and inform decision-making.

MICHAEL BASTEDO, PhD, is the Marvin W. Peterson Collegiate Professor of Education in the Center for the Study of Higher and Postsecondary Education at the University of Michigan. He has scholarly interests in the organizational decision-making of public higher education, particularly college

admissions, enrollment management, stratification, and rankings.

JOHN GONZALEZ, PhD, is the Director of Institutional Research at the Rackham Graduate School at the University of Michigan. He leads the Michigan Doctoral Experience Study (MDES), a longitudinal exploration of the doctoral student experience, and his work focuses on understanding how interventions, policies, and disciplinary context shape the experience of graduate students.

ALLYSON FLASTER, PhD, is a senior research analyst at the University of Michigan. Her recent work focuses on equity in graduate admissions, measuring the impacts of liberal education, and supporting data sharing among the educational research community.

YIPING BAI, MA, is a doctoral student in the Center for the Study of Higher and Postsecondary Education at the University of Michigan. Her research interests involve studying the education inequality among race/ethnicity, socioeconomic status, and urbanicity, as well as evaluating the effects of education policies along these lines, to more equitably serve different student populations.

Manuscript received September 9, 2024

Revised July 7, 2025

Accepted September 8, 2025