Who Benefits From Multiple Choice(s)?: The Equilibrium Impacts of Test-Optional College Admissions

Jack Kelly*

Senior Essay, Yale University Dept. of Economics, April 2022

Abstract

A rapidly increasing number of American colleges and universities have implemented test-optional admissions policies that allow applicants to withhold their SAT/ ACT scores, occasioning a lively public dialogue but little formal scholarship. This paper estimates the equilibrium effects of such policy by asking which students are more likely to get in under test-optionality than when required to submit scores (what I term "full disclosure" admissions). I develop a general framework that can be used to simulate the effect of a switch to full disclosure given data from a test-optional status quo, and apply this to a proprietary data set of all applicants from a test-optional college between 2017 and 2019. My simulation results suggest that a switch to full-disclosure would have little impact on the share of admitted students who are underrepresented minorities and the share who are non-submitters. My findings indicate that full disclosure has minimal effects on these quantities both because (i) few students see their admissions outcomes change and (ii) among those whose outcomes do change, those students who are benefitted and those who are harmed by the switch to full disclosure are demographically similar. I further analyze different models of behavior under test-optionality to examine how universities may behave in the absence of scores.

*B.A. Candidate, Economics & Mathematics, e-mail: jack.kelly@yale.edu. I would like to thank my advisor, Professor Jason Abaluck, for constant support, insight, and patience throughout the process of writing this senior essay. I never could have navigated the curcuitous nature of either the research process or the SOM stairwells without his guidance. I also benefitted greatly from the assistance of the senior essay teaching fellow Natalia Drozdoff, and from conversations regarding this paper with Professors Mira Frick and Eduardo Dávila. Professor Dávila also was kind enough to employ me as a a research assistant of his for over a year a half, a period during which he taught me an immense amount, ranging from the beautiful subtleties of normative economics to how to pronounce "dplyr." I thank the Economics Department's Tobin and Scarf programs for making such research opportunities possible. I also wish to thank the Mellon Senior Research Grant for financially supporting this project. This paper uses proprietary data from a test-optional college; I would like to extend my deep gratitude to the admissions department at this college for agreeing to share their anonymized data with me and for replying to every question I emailed them with inspiring speed and thoroughness. My friends and family have provided vital emotional support throughout the essay-writing process. I particularly wish to thank my mother, Madelyn Kelly, who, along with constant encouragement, has texted me every single article about test-optional admissions that has appeared in a major news outlet since I told her about the nature of my project. Had I chosen any other topic, I know she would have done the same. Any errors are my own.

Contents

"[Going test-optional] is about doing the RIGHT thing... [w]hich is helping students and families of all backgrounds better understand and navigate this process and about bringing students with intellectual promise (no matter their background) to UChicago (and making sure they succeed here too!)." – James G. Nondorf, Vice President for Enrollment and Student Advancement and Dean of College Admissions and Financial Aid, the University of Chicago, June 14, 2018. [1]

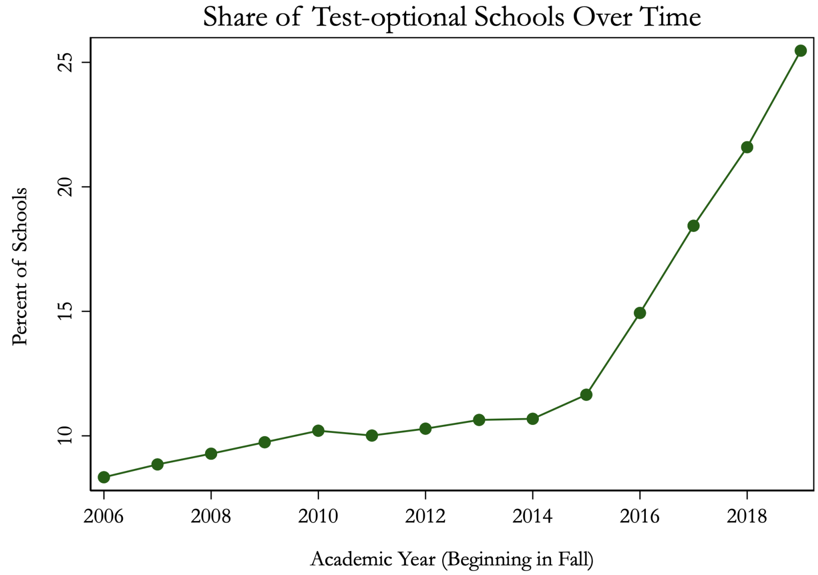
"We are reinstating our [SAT/ACT] requirement, rather than adopting a more flexible policy, to be transparent and equitable in our expectations. Our concern is that, without the compelling clarity of a requirement, some well-prepared applicants won't take the tests, and we won't have enough information to be confident in their academic readiness. Again, our research suggests this is most true for our most disadvantaged applicants, whose other educational opportunities have been most disrupted by the effects of the pandemic when they apply. We believe it will be more equitable if we require all applicants who take the tests to disclose their scores." – Stu Schmill, Dean of Admissions and Student Financial Services, Massachusetts Institute of Technology, March 28th, 2022. [2]

1 Introduction

After playing a key role in American undergraduate admissions for almost a century, standardized tests such as the Scholastic Aptitute Test (SAT) and American College Test (ACT) are rapidly becoming an elective component of college applications. A growing number of schools have adopted "test-optional" policies, whereby applicants are permitted, but not required, to submit standardized test scores for consideration. As Figure 1 documents below, the upwards trend in test-optionality from 2006 to 2019 has been dramatic, and the COVID-19 pandemic has led to even more institutions becoming, at least temporarily, test-optional, including the entire Ivy League and many flagship state schools. This increase has coincided with a spirited public dialogue speculating on the potential costs and benefits of test-optionality. Schools adopting the policies, as well as other advocates, tend to depict them as diversity-enhancing, as captured in the first quote above. These arguments tend to focus on the costs of testing and sending scores as well as well-documented disparities in standardized test scores across lines of race, ethnicity, and class. By contrast, as encapsulated in the second quote, critics of test-optionality have argued that the policy may disfavor marginalized applicants, for instance by shifting weight towards other metrics that are even more biased, or by leading applicants who are in fact better off submitting to withhold their scores. Implicitly, this discourse suggests that a key prerequisite to many people's normative assessment of test-optional policies is answering the question of who stands to gain and who stands to lose from their adoption relative to the historically typical policy of requiring all students to submit scores – which I will call "full disclosure."

[1] See here.
[2] See here.

Figure 1: The Rising Share of Test-Optional Schools, 2006–2019



Source: IPEDS, individual Common Data Sets, university websites. See Appendix D for details on classifying test-optional schools.

Taking the statements of many test-optional colleges at face value, the answer to the question posed in the preceding paragraph would be that *all* applicants at least weakly benefit from the policy. Consider how Roger Williams University, a test-optional institution, advises prospective applicants:

> "If you feel that your SAT and/or ACT scores accurately reflect your academic achievements and potential you may wish to submit your scores. Conversely, if you feel that your transcript and other involvements and achievements are a better reflection of your ability and potential, and that your test scores detract from that, you may choose not to submit them."[3]

Quotes such as this one suggest a staightforward appeal of test-optional policies: by turning a constraint into a new choice variable, all applicants would appear to gain. But there are at least two reasons why such an intuition would be suspect in this environment. The first is that applicants' decisions of whether or not to submit may be interdependent, and as such impose externalities on others. For instance, if low-scoring applicants choose not to submit while high-scoring applicants opt to send scores, colleges may (correctly) infer the non-submitters to have low scores. The second reason is that if colleges are constrained in how many students they can admit (say, by the number of rooms or class slots they have available), the net impact of applicants' submission decisions may force colleges to tighten or loosen the threshold above which they admit applicants: for instance, if relative to a full disclosure equilibrium, all submitters are seen as equally desirable and all non-submitters are seen as strictly more desirable, a college whose capacity constraint binds would need to raise its admissions standards to maintain the same number of admits.

In other words, to fully answer the question of who stands to gain or lose from test-optional admissions, it is important to estimate the *equilibrium effects* of test-optional policies. Facially, this might seem a daunting task: the introduction of a further

[3] See here.

choice variable to students' application decisions opens a rich set of channels that are challenging both to theoretically model and to empirically test. In this paper, using data from a test-optional college, I show that it is possible to sidestep many of these questions by in a sense inverting the question: starting with a test-optional status quo, and simulating who would be admitted if the university switched to full disclosure. The key asymmetry that motivates this empirical approach is the following: when starting from full disclosure and simulating the effect of a switch to test-optionality, it is necessary to model both (i) students' "submission function" regarding whether or not to disclose scores and (ii) colleges' "admissions function" for the non-submitters, which there is little a priori reason to think would be identical to their admissions function for those students under full disclosure. Further complicating matters, since students would plausibly form their (non-)submission decision based on the perceived impact on their probability of admission, and colleges would rationally form inferences on the scores of non-submitters as a function of who chooses not to submit, these two functions are codetermined in equilibrium.

By contrast, if one starts with test-optional data and simulates the effect of a switch to full disclosure, (i) every student must submit, so the decision-making function is irrelevant, and (ii) under plausible assumptions, colleges will simply evaluate *all* applicants according to the function they used for the *submitters* under the test-optional regime. Given suitable further assumptions on the distribution of observable and unobservables for submitters and non-submitters, I show one can then estimate admissions probabilities under full disclosure using the estimated coefficients of the submitters' status-quo admissions function. In particular, one can analyze the impacts under either the assumption that (i) colleges are capacity-constrained and will admit the same number of applicants under either admissions regime or (ii) colleges are unconstrained in how many students they can admit and will simply admit all applicants above a certain quality threshold under either regime. While this approach does require the strong assumption that the same individuals apply to the college under either admissions policy, it provides a baseline framework to motivate further work.

I operationalize this methodology on a proprietary data set comprising all applicants to a test-optional college over the last several years, restricting my attention to before the onset of the COVID-19 pandemic. My results suggest that the overall impacts of a switch to full-disclosure would change little on aggregate: the estimated percentage of admitted students who submitted scores under test-optionality increases by roughly one to two percentage points under full disclosure, from a status-quo average (across years) of roughly 73.6% to a simulated average of approximately 75.9%. Moreover, the difference in these percentages is not always statistically significant despite relatively tight confidence intervals. The change in the share of students predicted to be admitted who are underrepresented minorities (URM) is never significant in any of the years in my sample, and even if the point estimates are taken at face value, the magnitude of the changes are consistently small. These aggregate impacts mask subtler displacement effects: among the applicants predicted to be rejected under full disclosure who are admitted in the status quo, non-submitters are indeed overrepresented. However, non-submitters are also well-represented among the converse group: students predicted to be admitted under full disclosure who are rejected in the status quo, muting the aggregate effects. I offer evidence to suggest that a further reason for the minute effects is that the marginal effects of SAT scores on admission do not appear to be very large in magnitude relative to the effects of other dimensions along which non-submitters are internally heterogeneous, such as high school grade point average (GPA) – meaning that under both test-optionality and full disclosure, those factors are more determinate of which applicants are admitted. Meanwhile imposing a capacity constraint such that the college admits the same number of students under either policy changes little relative to allowing them to maintain a constant quality threshold above which they admit applicants, suggesting that general equilibrium effects, while a priori important to consider, are quantitatively minor in my setting.

While my simulation approach allows me to avoid taking a stance on how exactly non-submitters are treated in the status quo at test-optional schools, one may nevertheless be curious regarding these individual mechanisms, as it may inform the external validity of my findings. I show that a very basic model would predict full unraveling, where all but the lowest-scoring applicants still submit scores under test-optionality. I then consider three phenomena that could explain why empirically, some applicants choose not to submit: reweighting of the admissions function towards other metrics for non-submitters, "cursed" equilibria where applicants or college neglect the informational content of non-submitters' choices/covariates, and rankings-sensitivity where colleges care differentially about the scores of submitting students due to the effect on college rankings. I provide empirical tests that suggest that at least in certain strong forms, the former two phenomena do not appear to hold in the data, while I find suggestive evidence for the presence of rankings-maximization.

The rest of this paper is organized as followed: Section 2 discusses related literature; Section 3 describes my data and provides institutional background on the college; Section 4 presents my framework for simulating admissions under full-disclosure and lays out the results of this exercise; Section 5 offers models of admissions under test-optionality and provides empirical tests of some of these models, and Section 6 briefly concludes.

2 Related Literature

My work relates to at least four distinct literatures.

First and most directly, my paper speaks to the small but extant literature on test-optional admissions in American colleges and universities. Among this scarce literature, my paper fits into an even smaller subset that obtains and employs microdata from test-optional college, comprising the contributions of Robinson and Monks (2005), Conlin, Dickert-Conlin, and Chapman (2013), and Conlin and Dickert-Conlin (2017). These latter two papers in particular, which focus respectively, on how college rankings interact with test-optionality and on the potential for colleges to incorrectly impute scores of non-submitters, address issues similar to the potential mechanisms of rankings maximization and cursed equilibria discussed in section 5, although my empirical approach departs from these authors'. In particular, I show in section 5.3.3 that their approach to testing for rankings maximization is difficult to theoretically justify, as it would predict full unraveling, and I instead propose a modification of their approach that accounts for universities' imputations of students' scores. More generally, my paper differs from the existing literature using microdata from test-optional colleges in that my main focus is on simulating the equilibrium impacts of a switch (back) to mandatory disclosure of test scores, as I explore in section 4. By contrast, the individual mechanisms examined in these papers speak more to the question of whether *individual* applicants or adimssions department are acting in their best interest within a test-optional regime. While these channels are interesting in their own right, they are neither necessary nor sufficient to identify the equilibrium impacts of the policy relative to the full-disclosure alternative, which is presumably the relevant question for schools and for those interested in welfare analysis.

Another subset of the literature on test-optional admissions that does focus more on equilibrium effects has examined the impacts on aggregate quantities at the college level, such as admissions and yield rates, the demographic composition of matriculants, and average test scores of matriculants. These papers, which include Saboe and Terrizzi (2019), Belasco, Rosinger, and Hearn (2015), Sweitzer, Blalock, and Sharma (2018), and Bennett (2022), primarily use data from the Integrated Postsecondary Education Data System (IPEDS) and employ staggered-timing difference-in-difference designs to estimate the causal impact of test-optional admissions. With the exception of the last article (which uses a slightly more recent data set and different empirical specification), these papers have tended to find insignificant effects on demographic diversity from switching to test-optionality, but significant reductions in the admissions rate and increases in the yield rate.

Relative to my microdata-based approach these papers offer both advantages and disadvantages. In terms of advantages, by examining a larger set of test-optional colleges (sometimes all of them) and using many non-test-optional colleges as the control group, these papers can provide more generalizable findings than my results, which are arguably relevant only to the college for which I have data or very similar ones. Secondly, by dint of having access to data both before and after schools go test-optional, these papers can capture channels that may be difficult to simulate counterfactually with a set of test-optional data only, for instance, potential impacts on students' choices to apply to certain colleges in the first place now that they need not submit their scores.

On the other hand, the aggregate college-level data these papers use are rather limited in scope, constraining the inferences that can be drawn: for one, most IPEDS statistics (e.g. demographics, average scores) are taken among matriculants, not admits, making it difficult to disentangle the effects of test-optionality on admissions from those on enrollment conditional on admission. Additionally, the IPEDS does not require test-optional schools to fill out the portion of the survey that concerns standardizing testing – meaning key data points like the share of applicants submitting scores and the quartiles of SAT/ACT scores among matriculants are missing for the majority of test-optional schools, and the subsample for which they are available may not be random. Beyond the data constraints, these papers may also be susceptible to some of the challenges that frequently besiege differences-in-differences designs, such as the validity of the parallel trends assumption; while many of the papers do test for pre-trends, the relatively low number of treated units in the sample period suggests that the concern pointed out by Roth (2022) might be acute in this setting: that insignificant pre-trends need not necessarily be indicative of underlying parallel trends, but may simply be a consequence of low power. Moreover, the strategy undertaken by these papers of estimating causal effects in a staggered-timing differences-in-differences design through a two-way fixed effect linear regression has been shown to be inconsistent in the presence of heterogeneous treatment effects (see, for instance, Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Goodman-Bacon (2021)), and the impacts of test-optionality presumably do vary across colleges. Lastly, even if these papers meet the requisite conditions for consistently estimating causal effects via differences-in-differences, the confidence intervals they document are large, likely owing again to the small number of treated units: it is thus difficult to discern whether the null effects on many variables of interest (e.g., share of minority students) owe to the policy genuinely having little impact, or to low statistical power. By contrast, the null results I document in section 4 are relatively tight, allowing me to rule out considerably sized effects.

A second relevant literature focuses on jointly modeling students' application decisions and colleges' admissions decisions; contributions include Kapor (2020), Bleemer (2021), Blair and Smetters (2021), and Chade, Lewis, and Smith (2014). The majority of these models expend considerable energy on students' initial application decisions and the construction of college "portfolios"; this question, while likely influenced by test-optionality, is unfortunately not one my data are well-suited to address, and thus I focus on the impacts on admissions probabilities taking application choices as given. Nevertheless, many of the insights of these models guide my work: for instance, the model in Section 4 shares the framework in Kapor (2020) of using ratings functions and admissions threshold to represent admissions decisions. Meanwhile, the model of Blair and Smetters (2021), which highlights the role of maximizing perceived selectivity in college admissions, helps motivate the investigation of "rankings maximization" in section 5.2.3. Furthermore, my findings connect intriguingly to the subset of this literature (for instance, Kapor (2020) and Bleemer (2021)) that has focused on the effect of automatic admissions programs such as Texas and California's "top percent" policies, whereby students in the upper percentiles of the GPA distribution at their high schools are automatically admitted to top state universities. These papers have tended to find larger effects than mine in terms of the changes in the demographic composition of the admitted class compared to the simulated class that would have obtained under standard admissions. While the top-percent setting is facially similar to mine in that it too makes test-scores less relevant for the admissions functions of some students, a key difference is that top-percent complying schools are forced to admit all who meet the GPA threshold, whereas test-optional schools are free to choose a different admissions function for non-submitters.

Third, my work connects at a more abstract level to the long line of research on adverse selection stemming from the seminal contribution of Akerlof (1970). More specifically, the baseline "unraveling" result in Section 5.1 can be seen as an analogue of the theory of costless signaling in Grossman (1981): if the costs of sending a signal is assumed to be zero and lying is impossible, disclosure from one party of anything less than full information leads rational counterparties to form pessimistic inferences. Meanwhile, the exploration of possibly "cursed" equilibria wherein either applicants and/or admissions officers underestimate the degree to which score non-submission is informative of a student's underlying score draws directly from the concept coined and formalized by Eyster and Rabin (2005).

The exploration of potential misoptimizations such as a cursed equilibrium also connects my paper to a fourth literature, which concerns the scope for irrationality in the college admissions process. Pallais (2015) examines the effect of a reform that added an extra free ACT score report to college-applying seniors and finds that the elasticity of applications with respect to this relatively minute change in costs is very high, a fact that she argues is difficult to rationalize given the high returns to attending even a marginally better college. [4] Hoxby and Avery (2012) document low rates of applications to top colleges among high-achieving low-income applicants, despite the fact that these applicants have high probabilities of admission, and, conditional on admission, would likely face lower out-of-pocket costs than at lower-ranked schools due to extensive financial aid, a finding that in principle could owe to behavioral biases, informational barriers, or heterogeneous preferences. Dynarski et al. (2021) offer support for the behavioral hypothesis by examining the effect of simplifying the financial aid application process at the University of Michigan via a randomized experiment, promising some low-income students ex ante that if admitted, they will qualify for the amount of aid they would receive in expectation ex post; this intervention significantly boosted applications and enrollment rates among the treatment group.

3 Institutional Details, Data Description, and Summary Statistics

3.1 Institutional Details

My data comprise all prospective first-year applicants to a single test-optional college over the course of four years (2017–2020). While I will not state precisely when the college went test-optional to help preserve its anonymity, the policy had been in place for over ten years from the beginning of my data set. The college is private and not-for-profit. Over the entire sample period, the college was on the Common Application, a platform whereby students can enter the same information (e.g. demographic details, extracurriculars, essays) for some portions of many different colleges' application forms. Prospective first-year students may apply to the college in one of four rounds: Early Action (EA), Early Decision 1 (ED1), Early Decision II (ED2), and Regular Decision (RD). The first two options, EA and ED1, typically have application deadlines in early November, with students hearing back regarding their status in early January and early December respectively. The ED2 and RD application deadline is usually in mid-January, with students notified of their admissions decision around late February for ED2 and early April for RD. The ED1 and ED2 options differ from EA and RD in that the former channels are binding: if a student is admitted via these pathways, they

[4] Alternatively, Chade, Lewis, and Smith (2014)argue that a high elasticity to the marginal cost of applying could be rational if the admissions probabilities across colleges are sufficiently uncorrelated; for instance, in the extreme case in which all college admissions are i.i.d. with common admissions probability , the probabilitiy of not getting into N other schools forms a geometric sequence in N, easily diluting even large returns to education.

must matriculate to the college. EA and RD students have until early May to ultimately decide whether to matriculate.

3.2 Preliminary Data Cleaning and Baseline Sample

In the raw data given to me by agreement with the college, I observe data for all 44,846 students who applied for the first-year classes beginning in Fall 2018 through Fall 2021. I drop from the sample the 11,129 students who applied for the first-year class matriculating in 2021, the reason being that these students applied during the COVID-19 pandemic, during which the treatment of test-optional students may have been rather different; for instance, the college would have presumably been more likely to assume that non-submitters could not (safely) access a testing site rather than that they had taken the SAT/ACT and obtained low scores they wish not to share. I also drop all 74 recruited athletes from the sample; these students tend to face a functionally distinct admissions process than other students. I drop two students who are listed as enrolled but for whom the "admitted" variable is equal to zero (indicating rejection) and one student for whom the SAT (equivalent) score is reported as above 1600; these three observations are presumably errors in data entry. This cleaning results in a data set of 33,253 observations, each corresponding to a unique applicant.

3.3 Available Variables

I observe the following covariates for almost all applicants: year of intended matriculation, gender, race/ethnicity, IPEDS racial/ethnic classification, state and country where high school is located, high school GPA (standardized to a common 4-point scale), curriculum score (the admissions' office's assessment of the difficulty of a student's high school curriculum, on a 10 point scale), intended major, student application round (EA, ED1, ED2, or RD) and whether or not a student submitted test scores. I also observe the outcome variables of whether a student is admitted, and whether a student ultimately enrolled.

I observe test scores for all submitters and for some non-submitters (the subsample of non-submitters who have data available is discussed in detail in section 3.5 below). Test scores take the form of SAT math and reading/writing scores or ACT composite scores; to standardize the analysis and because the SAT admits a finer gradation of scores (400–1600 with 10 point intervals rather than 1–36 with 1 point intervals) , I form SAT total score as the sum of the reading/writing and math scores, and convert all ACT composite scores to SAT total scores using the official concordance table. [5] If a student submitted both the SAT and the ACT, I take their score to be the maximum of the SAT and the SAT-equivalent-converted ACT.

Private correspondence with the admissions office confirmed that the college does not consider either the SAT Subject Tests or Advanced Placement (AP) tests when making admissions decisions. The major potentially admissions-relevant variables I do not observe are thus students' "intangible" components: essays, extracurriculars, and letters of recommendation, as well as whether or not an applicant is a "legacy" (the child of one or more college alums).

3.4 Classifying Test-Optional Applicants

There are two columns in the data that can suggest a student is applying under test-optionality. First, students may indicate on their application directly whether or not they wish to have scores considered; separately, students can choose to send their scores by listing them in the self-reported score section of the Common Application, having their high school print it on their transcript, or directly sending scores via the College Board or ACT Corporation to the admissions office. From private correspondence with the admissions office, I have confirmed that any applicant who indicates on the application that they would not like their scores

[5] The table can be found here

considered will be treated as test-optional even if they do self-report or send their scores to the college. There are, however, also applicants who answered that they would like to have their scores considered, but do not have visible scores in the data. The admissions office confirmed to me that these applicants never reported scores via any of the mechanisms listed above, and were simply treated as test-optional; I thus recode all such applicants as test-optional. In principle, a potential mechanism for misclassification here is if a student indicated they would like scores considered when they applied, did not send scores before the deadline (so they were treated as test-optional), and was then admitted and sent scores to the school after admission (the reasons for which are discussed below). Such students would be classified as not test-optional in my dataset even though they were treated as such by the college at the time of application, however, these are presumably a very small subset of the total number of students.

3.5 Test-Score Availability Among Test-Optional Applicants

While all students defined as submitters have test scores available (partially by construction, as described above), the same cannot be said for non-submitters. Table 1 below documents the total number and relative frequency of non-submitters with and without scores across three outcome groups: the rejected students, admitted but not enrolled students, and enrolled students. There are two channels through which scores can be visible for applicants who are test-optional. The first is the aforementioned one: some students submit their scores either officially or via self-reporting on their application, but also tick the test-optional box indicating they would not like to have scores considered. The second is that admitted students are asked to send their scores to the college for advising; private correspondence with the admissions office has confirmed that students receive two emails asking them to submit but are not bound to doing so, hence the incomplete compliance documented below. Of course, for the admitted or enrolled students, we cannot discern which of the two mechanisms explains the availability of the data, while the rejected non-submitters with visible scores clearly must have reported or sent their scores with their application.

Table 1: Test-score Data Availability Among Non-submitters

	Rejected	Admitted, Not Enrolled	Enrolled	Total
	Total No.	Total No.	Total No.	Total No.
Data Availability	(Relative %)	(Relative %)	(Relative %)	(Relative %)
Scores Not Available	5,755	2,516	560	8,831
	(81.40)	(79.70)	(46.24)	(77.21)
Scores Available	1,315	641	651	2,607
	(18.60)	(20.30)	(53.76)	(22.79)

3.6 Descriptive Statistics

3.6.1 Frequency of Non-Submission

A natural starting point is to wonder how many students in fact exercise the option to not submit their scores. Panel A of Figure 2.2. below shows the rates of (non)-submission over the course of the three years I analyze. The fraction of applicants who do not submit scores hovers around 1/3 throughout the entire time period, with modest increases in the share not submitting year over year. Panel B of the same figure stratifies submission rates by the application round; interestingly, the non-submission rates appear markedly higher in the binding application rounds (ED1 and ED2) than in the non-binding rounds (EA and RA), although the sample size for the binding rounds is small, even aggregated across years.

Table 2: Share Submitting Scores by Year and Application Round

Panel A: Submission Rates By Year

	Year			
	2018	2019	2020	Total
Submitted Scores	Total No. (Percent)	Total No. (Percent)	Total No. (Percent)	Total No. (Percent)
N	3,679	3,899	3,860	11, 438
	(32.66)	(34.41)	(36.23)	(34.40)
Y	7,587	7,433	6,795	21,815
	(67.34)	(65.59)	(63.77)	(65.60)
Total	11,266	11,332	10,655	33,253

Panel B: Submission Rates By Application Rounds

	Application Round				
	EA	ED	ED2	RA	Total
Submitted Scores	Frequency (Percent)	Frequency Percent)	Frequency Percent)	Frequency (Percent)	Frequency (Percent)
N	5,519	439	214	5,266	11,438
	(28.64)	(53.60)	(52.84)	(41.27)	(34.40)
Y	13,751	380	191	7,493	21,815
	(71.36)	(46.40)	(47.16)	(58.73)	(65.60)
Total	19,270	819	405	12,759	33,253

3.6.2 Application Outcomes and Submission Choices

Given this paper's focus on the impact of test-optionality on admission, it is useful to simply characterize how the admissions rates of submitters compare to those of non-submitters, cognizant that such differences need not be causal. Panel A of Figure 2.3 below characterizes the respective shares of submitters and non-submitters who are admitted by the college, showing the share of admits to be much higher for the submitters. Panel B characterizes, among the respective subpopulations who are admitted, the shares of submitters and non-submitters who ultimately enroll in the college. While the share enrolling (the "yield rate") is low for both groups, it is markedly higher among non-submitters, a fact that could be consistent either with them benefitting from test-optionality (and thus faring worse at other non-test optional colleges where they must submit scores) or simply with the non-submitters faring worse on other admissions-relevant dimensions (like GPA), lowering the probability they get into higher-ranked colleges independent of their submission choices.

Table 3: Share admitted and enrolling by admit status

Panel A: Admit Rates, Submitters vs. Non-Submitters

	Submitted Scores		
	N	Y	Total
	Frequency	Frequency	Frequency
Admitted	(Percent)	Percent)	(Percent)
N	7,070	9,682	16,752
	(61.81)	(44.38)	(50.38)
Y	4,368	12,133	16,501
	(38.19)	(55.62)	(49.62)
Total	11,438	21,815	33,253

Panel B: Enrollment Rates, Admitted Submitters vs. Admitted Non-Submitters

	Submitted Scores		
	N	Y	Total
	Frequency	Frequency	Frequency
Enrolled	(Percent)	Percent)	(Percent)
N	3,157	10,242	13,399
	(72.28)	(84.41)	(81.20)
Y	1,211	1,891	3,102
	(27.72)	(15.59)	(18.80)
Total	4,368	12,133	16,501

3.6.3 Academic Performance Variables and Submission Choices

Given the lower admissions rates among non-submitters, one may wonder how the non-submitters compare to their score-submitting counterparts. Panel A of Figure 2.4 below shows that the mean standardized test score, high school GPA, and curriculum score of the non-submitters are all lower than the respective means among submitters. Since the subsample of non-submitters with test scores contains a relatively larger share of admitted non-submitters relative to the overall set of non-submitters, it is perhaps more informative to investigate the means separately for the admitted and rejected students among both submitters and non-submitters, as shown in Panels B and C. Interestingly, the average test scores among the admitted *non-submitters* lie significantly below those of the rejected *submitters*. Given the extent of this deficit, one might expect that the admitted non-submitters would be those who "compensate" for low (imputed) test scores with high GPAs and curriculum scores. In fact, while the average GPA and curriculum score of admitted non-submitters exceeds those of the rejected non-submitters, they are nevertheless lower than those of the admitted submitters.

Table 4: Academic variables by submission choice and application outcome

Panel A: Academic Indicators, Submitters vs. Non-Submitters

	Submitted Scores		
	N	Y	Total
Academic Metric	Mean (Std. error)	Mean (Std. error)	Mean (Std. error)
test_score	1110.8	1293.7	N/A
	(2.557)	(0.798)	
GPA	3.239	3.412	3.353
	(0.00428)	(0.00273)	(0.00236)
curriculum_score	7.235	7.881	7.659
	(0.0117)	(0.00844)	(0.00706)

Panel B: Academic Indicators, Admitted Non-Submitters vs. Rejected Non-Submitters

	Admitted		
	N	Y	Total
Academic Metric	Mean (Std. error)	Mean (Std. error)	Mean (Std. error)
test_score	1067.3	1155.1	N/A
	(3.820)	(2.916)	
GPA	3.051	3.539	3.239
	(0.00533)	(0.00419)	(0.00428)
curriculum_score	6.804	7.923	7.235
	(0.0135)	(0.0170)	(0.0117)

Panel C: Academic Indicators, Admitted Submitters vs. Rejected Submitters

	Admitted		
	N	Y	Total
Academic Metric	Mean (Std. error)	Mean (Std. error)	Mean (Std. error)
test_score	1239.7	1336.7	1293.7
	(1.194)	(0.898)	(0.798)
GPA	3.136	3.630	3.412
	(0.00393)	(0.00232)	(0.00273)
curriculum_score	7.169	8.446	7.881
	(0.0110)	(0.00970)	(0.00844)

3.6.4 Student Demographics and Submission Choices

Given the focus within the public discourse on test-optionality on potential benefits or harms for historically underrepresented demographic groups, it is suggestive to examine how submission rates vary by race/ethnicity[6] and gender. As documented in Table 16 in Appendix Appendix A, historically underrepresented racial and ethnic groups appear to be overrepresented among the non-submitters relative to the submitters, and women appear overrepresented in the non-submitters (the application appears

[6] I say "race/ethnicity" because the relevant variable in the data primarily reports race but also contains an option for Hispanic students of any race

to only have an option to indicate a binary gender identity).

4 Simulating Admissions Under Full Disclosure

In this section, I address the core question of this paper: *how many applicants would have the same admissions outcomes if all students were required to submit scores*? This task – comparing a test-optional equilibrium to one of full disclosure – is clearly relevant from a policy perspective: the answer may directly inform, for instance, whether a university decides to go (or stay) test-optional. I first characterize a general model of admissions and then lay out the key econometric assumptions under which one can estimate the impact of a switch from test-optionality to full disclosure using data from the test-optional status quo. I then apply this procedure to my data and present my main results.

4.1 A General Model of Test-Optional Admissions

The environment is populated by one college and a unit mass of student indexed by $i \in \mathcal{I}$. Each applicant is endowed with a test score SAT_i, which is a bounded random variable normalized to lie in $[0, 1]$ with density $f(\cdot)$, and a vector $\mathbf{X}_i \in \mathbb{R}^n_{++}$ of non-test score covariates observable to colleges; I will often collect these attributes in the vector $\mathbf{Z}_i = \{\text{SAT}_i, \mathbf{X}_i\}$ for brevity. [7]

Students who have taken the SAT can choose to submit scores or to withhold them, with no cost of submitting and no possibility of submitting a score that differs from their true one, i.e., they cannot lie. A priori, I take no further stance on the student's decision-making function on whether or not to submit, other than to assume they play a pure strategy, such that for each student we can define the function $S_i = \mathbb{I}\{\text{student } i \text{ submitted test scores}\}$.

Students with $S_i = 1$ (henceforth "submitters") are given a rating by the admissions committee determined by a function $g(\mathbf{Z}_i)$ that reflects the payoff of the student being admitted from the point of view of the college. Non-submitters are evaluated according to a (possibly different) rating function $h(\tilde{\text{SAT}}(\mathbf{X}_i), \mathbf{X}_i)$. The function $\tilde{\text{SAT}}(\mathbf{X}_i)$ denotes the value of students' imputed test scores; a priori, this need not be the conditional expectation of non-submitters' scores, but simply is a stand-in for the score that they are treated as if they have. This function can condition on the full vector of covariates and the choice not to submit, but must be the same for any two non-submitters with identical covariates. Admitting a student entails a constant marginal cost κ, and we assume that the college cannot admit all students because it has some capacity limit each year, C_t. I now characterize the core assumptions on the agents within our model.

College Preferences and Behavior

Assumption 1. *(The College's Problem)*

(i) The college's preferences for students are independent of which other students they admit.

(ii) At least one element of the covariate vector $\mathbf{X}_i$ is continuously distributed.

(iii) The functions g, h have gradients that are nonzero in every argument over their entire respect supports.

(iv) The college is "selective" in the language of Kapor (2020). In other words, its capacity constraint binds and, in particular, it admits the measure of students up to which they have capacity in that year.

Assumption essentially imposes that the college has a complete, independent ranking over applicant characteristics, which, combined with the continuity assumption in part (ii), implies that there exists a unique threshold score π such that to maximize

[7] I refer to scores as SAT scores throughout for simplicity, in reality of course the model is agnostic about whether students submit the SAT or ACT

the perceived quality of students, colleges admits students iff their rating $r(\text{SAT}_i, \mathbf{X}_i, S_i) = S_i g(\mathbf{Z}_i) + (1 - S_i)h(\tilde{\text{SAT}}(\mathbf{X}_i), \mathbf{X}_i)$ satisfies $r(\text{SAT}_i, \mathbf{X}_i, S_i) \geq \pi$ (since the functions g, h can always be adjusted to account for fixed preferences for submitters versus non submitters, it is without loss of generality to assume the same threshold is used).[8]

Part (iv) of Assumption 1 could fail in two ways. First, it would fail if the college is in fact not selective: selective here refers not to a given admissions rate but rather to whether the capacity constraint binds; a non-selective college admits applicants until the marginal cost of admission outweighs the marginal benefit, with the marginally accepted applicant occuring before all beds are filled. This portion of the assumption is empirically testable: if the college were non-selective, then the year of application should have no impact on the probability of admission conditional on all the observable applicant covariates, assuming that there are no unobservable applicant covariates correlated with the year of application. The data plainly rejects this; when running a probit regression with all covariates, we can reject the null that the coefficients (and the average partial effects) of the year fixed effects are 0 at the 1% level.

The second way the assumption could fail is if colleges are not in fact admitting a given number of *admits* but rather of expected enrolled students. This latter possibility is what Kapor (2020) assumes: in his model the capacity limit equals the total expected number of enrolled students. My differing assumption can be defended on at least four grounds. First, if enrollment is independent of applicant covariates conditional on admission, it is entirely without loss of generality to consider admits. Second, an admissions office may plausibly set a target for the number of admits in its own right rather than expected enrollment because colleges care about their selectivity (acceptance rate) as an input into their rankings, as explored by Blair and Smetters (2021). Thirdly, admissions offices may find it preferable to constrain their number of admits rather than expected enrollment because of the inherently stochastic nature of the enrollment decision from colleges' point of view and the plausibly asymmetric costs of over versus under enrollment: because colleges may be much worse off if more students than expected enroll (since this could create a shortage of resources) than if fewer than expected do, they may find it preferable to simply admit a number of students equivalent to the capacity constraint. Fourthly, as a purely empirical matter, the year-to-year percent changes in the number of admitted students are lower than the analogous changes in the number of enrolled students, although both are small (on the order of a few percentage points). While in principle this could be explained by the capacity constraint and the yield rate happening to comove in opposite directions, that would be far from the most parsimonious explanation.

Student Preferences and Behavior Because colleges decision rules become deterministic conditional on covariates in this environment, students must perceive the admissions process with some noise to avoid the process being fully revealing. In particular, I assume that students perceive the score functions accurately but believe the threshold to be $\pi + \varepsilon$, where ε is some symmetrically distributed noise, i.i.d. across applicants, with infinite support. This implies that students believe their admissions chances to be $F(r(\text{SAT}_i, \mathbf{X}_i, S_i) - \pi) \in (0, 1)$. The assumption that students do not know deterministically if they will be admitted can be defended on the grounds of introspection, and theoretically, as this would imply no students who are rejected ever apply. One might object to the less obvious assumption that the support of ε is infinite such that, for instance, even very strong applicants have only probabilistic knowledge of their admissions outcome ex ante. However, if this were the case, given even minute costs to submitting scores, we would expect a marked uptick in non-submission amongst very strong applicants, a pattern not borne out in the data.

A stronger assumption I impose on student behavior is the following:

[8] As Kapor (2020) shows, if we instead assume the college is not capacity constrained/its constraint does not bind, it is still optimal for colleges to set a unique threshold π; in particular the threshold would simply be the marginal cost κ

Assumption 2. *(The Students' Problem): Students only care about the submission decision only to the extent it impacts their probability of admission. Ties are broken in favor of submitting scores.*

Much of Assumption 2 may appear to be already implied by the assumption that submission is costless; however, it does rule out further motives for non-submission such as shame at having a lower test score than one aimed for, even if the student believes the college would look upon that score favorably, or students withholding/submitting scores because of the potential impact it may have on their class placements conditional on admission and enrollment.

With this assumption in place, we can then define an equilibrium for our environment.

Definition 1. Given a ratings function $r(\text{SAT}_i, \mathbf{X}_i, S_i)$, a **test-optional equilibrium** in year t is defined by

(i) The "imputation function" $\tilde{\text{SAT}}(\mathbf{X}_i)$ for non-submitters.

(ii) The "admissions function(s)" $r(\text{SAT}_i, \mathbf{X}_i, S_i) = S_i g(\mathbf{Z}_i) + (1 - S_i)h(\tilde{\text{SAT}}(\mathbf{X}_i), \mathbf{X}_i$

(iii) The "submission function" $S : \mathbf{Z}_i \to \{0, 1\}$ such that $S_i = 1$ iff $r(\text{SAT}_i, \mathbf{X}_i, 1) > r(\text{SAT}_i, \mathbf{X}_i, 0)$

(iv) A threshold π_t such that the measure of students in the set $\{i : r(\text{SAT}_i, \mathbf{X}_i, 1) > \pi\} = C_t$, the capacity constraint in that year.

4.2 Admissions under Full Disclosure

I next define the counterfactual object of interest: an admissions regime under which all applicants have to submit scores.

Definition 2. A **full-disclosure equilibrium** in year t is an equilibrium wherein students are required to submit scores and colleges admit an equivalent number of students as in the test-optional status quo in that year.

Note that under the assumption that college preferences over applicants are independent, the function $g(\cdot)$ that is used to evaluate submitters in the status quo is the same function that would be used to evaluate *all* applicants under full-disclosure. However, in order to simulate students' counterfactual behavior, more structure is needed. In particular, I impose the following:

Assumption 3. *The same set of students apply to the college under test-optional and full-disclosure admissions.*

Effectively, Assumption 3 entails that

(i) All status-quo non-submitters in fact took the SAT/ACT, and thus applying is feasible for them under full-disclosure and

(ii) Applying is desirable under full-disclosure if and only if it is desirable under test-optionality, for submitters and non-submitters alike.

Part (i) of Assumption 3 can be defended on the grounds that, notwithstanding the rise in test-optional schools, the majority of 4-year public and not-for-profit colleges and universities still required standardized tests throughout the time period of my analysis, including almost all state university systems. Given the reasonably selective nature of the institution for which I have data, few applicants would have likely been content to limit themselves only to test-optional schools across their entire college portfolio.

Part (ii) of Assumption 3 is much stronger. In reality, there are both costs to applying to a school in the first place and to submitting scores; these costs, combined with potential changes in the probability of admission for submitting versus not submitting, could generate an equilibrium in which some students only apply under test-optionality or full-disclosure. However, there are three constraints that make it difficult in practice to assess impacts on the initial choice to apply in my setting. First, it is theoretically challenging to account for simultaneous application, submission, and admissions choices, all of which may be codetermined in equilibrium. Second, at best my data could contain individuals who applied under test-optionality but may have not applied under full-disclosure, but any individuals who would have applied under full-disclosure but not test-optionality are necessarily not represented, since I only have data from years in which the college was test-optional. Thirdly, even assessing whether individuals who applied in the test-optional status quo would have applied under full-disclosure is empirically difficult, because the costs of applying and submitting scores are heterogeneous in ways that my data do not account for (for instance, some low-income applicants qualify for fee waivers, and the cost schedule for sending scores depends on how many other colleges a student has already sent them to). Therefore, I impose the assumption that the applicant pool is identical across the two policies throughout the paper, mindful of its limitations.

Based on our assumption that the college's capacity is constrained, our counterfactual simulation must account not only for non-submitters potentially facing different ratings under full disclosure but also for equilibrium adaptations necessary to maintain the same number of admits. Proposition 1 shows that this can be achieved simply through finding a new threshold for admission.

Proposition 1. *Under the setup in section assumptions 4 and 5, there exists a unique threshold π' such that the number of admitted students is maintained in full disclosure as under test-optionality, i.e.*

$$\mu(i : g(\mathbf{Z}_i) > \pi') = \mu(\{i : S_i = 1 \wedge g(\mathbf{Z}_i) > \pi\}) + \mu(\{i : S_i = 0 \wedge h(\tilde{SAT}(\mathbf{X}_i), \mathbf{X}_i) > \pi\}) = C$$

where μ is the (product) measure defining the joint distribution of $\mathbf{Z}_i$.

Proof. Since the gradient of $f(\cdot)$, is nonzero in each component, and since an element of $\mathbf{Z}_i$ is continuously distributed, the function $\psi(\pi') = \mu(i : g(\mathbf{Z}_i) > \pi')$ is continuous and strictly decreasing; moreover, $\psi(-\infty) \geq C$ (every student would be admitted), while $\psi(\infty) \leq C$ (no student would be), so by the Intermediate Value Theorem a unique equilibrium exists such that $\psi(\pi') = C$. □

4.3 Econometric Assumptions

In this section I lay out the core econometric assumptions for estimating a full-disclosure equilibrium. These are twofold: the assumptions necessary to infer the predicted rankings of non-submitters from data on the submitters, and the assumptions necessary to estimate the underlying test scores of those non-submitters for whom we do not have data. I lay out these assumptions in turn.

Assumption 4. *(Consistent Estimation of the Admissions Function)*

(i) The function determining the admissions score for a student with visible scores can be written as $\beta'\mathbf{Z}_i + \varepsilon_i$ where $\mathbf{Z}_i = \{SAT_i, \mathbf{X}_i\}$ is a vector of covariates visible to the econometrician, and $\varepsilon_i \sim \mathcal{N}(0, 1)$.

(ii) The error term ε_i in (1) is independent of the covariate vector $\mathbf{Z}_i$ and students' submission choice S_i.

Part (i) of Assumption 4 imposes that the admission function is linear (in parameters) and allows for standard estimation via a probit regression which implies that, assuming the error term to be in fact normal, further assuming it to be a standard normal is without loss of generality so long as $\mathbf{Z}_i$ contains a constant term. More substantial is part (ii) of Assumption 4. In addition to imposing independence of the regressors from the error term, we impose identical distribution of the error term for the submitters and non-submitters so that the predictions of admissions probabilities for non-submitters using coefficients from the regression on submitters will be consistent. While the assumption of independence of ε_i and $\mathbf{Z}_i$ may appear restrictive, it may be plausible given the rich set of conditioning variables: while, for example, scores in and of themselves would likely be correlated with unobservables such as essay strength, one may think that after further conditioning on GPA, curriculum score, and demographic variables, the strength of essays, extracurriculars, etc. are effectively random.

Of course, fitting predicted admissions probabilities using the coefficients from the regression on submitters will be challenged by the presence of non-submitters for whom test scores are missing. My approach to fill in the missing data is to first partition the non-submitters into groups that may be unobservably rather different from each other: the non-submitters who ultimately enrolled (E) , who were admitted but did not enroll (ANE) , and who were rejected (R). Let the function $D(j)$ denote whether non-submitting applicant j has scores visible in the data.

Assumption 5. *(Imputing Missing Scores)*

For each of the sets of nonsubmitters $J \in \{E, ANE, R\}$ test scores SAT_j can be written as $SAT_j = \gamma' \mathbf{X}_j + \nu_j$ for all $j \in J$ where

(i) $\nu_j | \mathbf{X}_j$ is independent of $D(j)$.

(ii) $\nu_j | \mathbf{X}_j$ is normally distributed with mean zero and variance $\sigma_{\nu_j | \mathbf{X}_j}$

Because our goal is simply to predict students' scores and not to perform causal inference on the parameter vector γ, it is not necessary to assume that the error term is uncorrelated with the regressors $\mathbf{X}_j$, simply that this pattern of correlation is independent of whether or not data is visible, as described in part (i) of Assumption 5. This assumption is thus weaker than assuming the non-submitters with scores visible for each of the sets {E, ANE, R} are a representative subsample of those without scores visible in those sets. Nevertheless, this assumption could be violated if, even after conditioning on all the covariates, non-submitters with relatively lower or higher scores or more likely to have scores visible: for instance, this could occur because higher scoring non-submitters are more likely to initially send scores and then change their minds and tick the test-optional box, or because higher scoring enrolled non-submitters are more likely to take up the college's request to submit scores for placement, foreseeing more favorable placement. I test robustness to this assumption by conducting alternative imputations in Appendix B that are relatively more "optimistic" and "pessimistic" (i.e., tending to impute higher or lower average scores for the missing observations); these different specifications change the results very little from the baseline.

4.4 Estimation Procedure

I turn now to discussing how to estimate a full disclosure equilibrium that maintains the same number of admits, as characterized in Proposition 1. From the point of view of the econometrician, I search for an equilibrium in which the number of students predicted to be admitted equals the actual number admitted in that year, where a student is predicted to be admitted if their estimated probability of admission within a full-disclosure equilibrium exceeds .5. Proposition 2 lays out the procedure for estimating such an equilibrium.

Proposition 2. *(Procedure for identifying the counterfactual admits)*

Under assumptions 1–5 the following approach consistently estimates the students predicted to be admitted in a given year under full disclosure:

1. Run a linear regression for each of the sets $J \in \{\text{E, ANE, R}\}$ of SAT_j on the vector of observables $\mathbf{X}_j$ for the subset of observations in J that have scores available

2. Use the estimated coefficients $\hat{\gamma}$ from (1) to predict $\hat{\text{SAT}}_j$ for those students in J who do not have scores available.

3. Estimate the admissions function $\beta' \mathbf{Z}_i$ by running a probit regression over the sample of students who submit their scores in that year.

4. Fit the predicted probabilities $\Phi(\hat{\beta}' \mathbf{Z}_i)$ where $\Phi(\cdot)$ is the normal cdf, and $\hat{\beta}'$ are the estimated coefficients from step (3).

5. Count the number of students N who are actually admitted in that year.

6. Set p equal to the Nth-highest threshold value of $\Phi(\hat{\beta}' \mathbf{Z}_i)$ among the applicants in that year, such that the number of observations with $\Phi(\hat{\beta}' \mathbf{Z}_i) \geq p$ is as close as possible to the observed number of admitted students in that year.

Proof. Appendix C. □

For my baseline specification, I impute the scores of non-submitters with missing SATs by regressing on GPA and curriculum score, as well as fixed effects for application round, gender, intended major, and year of application. The assumption that the function relating the covariates to test scores is unchanged year-to-year (up to a constant) is largely to economize on power given the low number of observations with scores in the relevant subpopulations. By contrast, I allow the admissions function to change all coefficients each year, and as such run a separately probit regression for each year of a student's admissions outcome on their test score as well as all the aforementioned covariates used for the test-score imputations step. Because the procedure described above introduces sampling uncertainty at a variety of junctures, I estimate all standard errors for the simulated moments via the bootstrap, running 1,000 replications over the entire procedure outlined in Proposition 2. The reported confidence intervals for all bootstrap estimated standard-errors are bias-corrected. A student is classified as an underrepresented minority (URM) if they are Black or African American, Hispanic of any race, Native American or Alaskan Native, or Native Hawaiian or Pacific Islander.

4.5 Results

4.5.1 Main Results

Table 5 shows the main results of the simulation procedure described above. For each of the years in my sample, I compare four empirical moments over the set of admitted students to their predicted analogues under full disclosure: the share of admitted students who submitted scores, the average test score and GPA of the admits, and the share who are underrepresented minorities. Note that the years refer to the year in which prospective applicants would enter the college, i.e. 2018 refers to applicants seeking to begin college in Fall 2018.

The results suggest that the impacts of a switch to full disclosure would be remarkably minimal in terms of the aggregate composition of the admitted class. While we can reject the null of equality for the means of the share submitting test scores under test optionality and full disclosure in 2019 and 2020, the magnitude of the change is small, on the order of 1 to 3 percentage points. The latter two years also show greater increases in the average SAT score and average GPA when switching from test optionality and full disclosure. The positive comovement in these indicators is somewhat surprising, as one may expect that the marginally admitted students under test-optionality would have low scores but higher GPAs to compensate. Lastly, the fourth row shows that the impact on the change in enrollment of underrepresented minority students from switching to full disclosure is negative but never significant.

Table 5: Simulated Composition of Admits under Full Disclosure

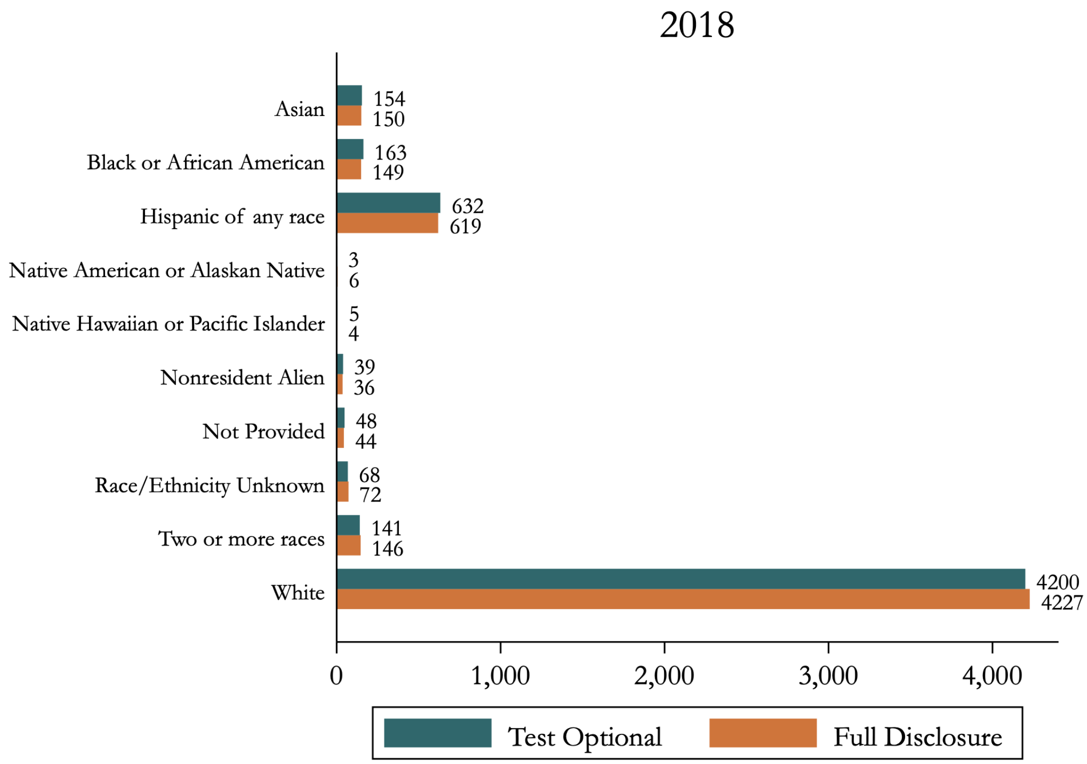
	2018		2019		2020	
	TO	FD	TO	FD	TO	FD
	Mean	Mean	Mean	Mean	Mean	Mean
	95% CI	95% CI	95% CI	95% CI	95% CI	95% CI
Share Submitting	75.21	76.66	75.42	77.79	70.16	73.21
	(74.06, 76.35)	(75.31, 77.91)	(74.26, 76.57)	(76.59, 79.21)	(68.98, 71.35)	(71.97, 74.54)
Avg SAT Score	1299.16	1304.64	1297.51	1305.34	1281.77	1293.17
	(1294.53, 1303.43)	(1299.58, 1308.44)	(1294.33, 1300.69)	(1300.80, 1310.16)	(1278.56, 1284.98)	(1288.96, 1297.02)
Avg HS GPA	3.60	3.62	3.62	3.64	3.60	3.64
	(3.59, 3.61)	(3.62, 3.63)	(3.61, 3.62)	(3.64, 3.65)	(3.59, 3.61)	(3.63, 3.65)
Share URM	14.71	14.27	16.32	15.25	15.04	14.00
	(13.77, 15.65)	(13.17, 15.44)	(15.33, 17.31)	(14.31, 16.71)	(14.11, 15.97)	(12.76, 15.19)

Notes: *For each year in my sample, Table 5 shows the estimated quantities of interest in the test-optional status quo (column "TO") and in the simulated full-disclosure equilibrium (column "FD") The average SAT score in the test-optional status quo is calculated using actual scores when available and imputed scores, following the procedure described in steps (1) and (2) of Proposition 2. The 95% confidence intervals for all simulated full-disclosure moments and for the average SAT score in the test-optional status quo are the bias-corrected CIs obstained via the bootstrap using 1,000 replications, run over the entire procedure outlined in Proposition 2. The confidence intervals for the non test-score moments in the test-optional status quo are the standard parametric ones for sample means.*

While Table 5 analyzes the impact on the share of admitted underrepresented minorities, one may be interested in a more granular demographic analysis. Figure 2 graphically compares, across all available racial/ethnic classifications, how the number of students admitted in the status quo compares to the number predicted to be admitted under full disclosure, showing the results to be minute across classifications. The graphs for 2019 and 2020 are very similar and can be found in Appendix A.

One may also wonder not just whether some applicants are predicted to be newly admitted or rejected under full disclosure but by "how much": in this vein, Table 6 provides a probabilistic analogue to the deterministic results in Table 5. Here, the average empirical probabilities of non-submitting students and URM students being admitted respectively are compared to their simulated analogues under full disclosure. This calculation entails one technical complication not present in the simulated composition moments in Table 5 that I briefly sketch out here. The threshold probability p found by the procedure described in Proposition 2 can be inverted to yield the change in the actual admissions threshold, $\pi' = \Phi^{-1}(p)$. For the students with test scores observable in the data, I can then straightforwardly estimate the probability of admission under the new threshold by calculating $\Phi(\hat{\beta}'\mathbf{Z}_i - \pi')$. However, for the students with imputed test scores, more care must be taken; this is because the estimation step introduces addition variance that affects the computation of the probability, since in a probit model the coefficients $\hat{\beta}'$ are only identified after normalizing the variance to 1. As shown in the Appendix, the correct probability for the imputed observations is $\Phi\left(\frac{\beta_0 \hat{\text{SAT}}_i + \beta'_{-0}\mathbf{X}_i - \pi'}{\sqrt{1+\beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i}}}\right)$ where $\rho_{\mathbf{X}_i} = \text{Cov}(\nu|\mathbf{X}_i, \varepsilon_i)$ This immediately introduces two ambiguities: how to model the conditional variance $\sigma^2_{\nu|\mathbf{X}_i}$, and how to adjudicate the size of the correlation between the conditional error terms of the score-prediction regression $\nu|\mathbf{X}_i$ and the probit error term ε_i. In the baseline results below, I make the simplest possible assumptions, taking $\nu|\mathbf{X}_i$ to be homoskedastic (so that I can simply estimate σ^2_ν from the residuals of the prediction regressions)

Figure 2: Number of Admits By Racial/Ethnic Classifications: Test Optional versus Full Disclosure



and ρ to be 0. The results in Table 6 reinforce those in Table 5: the average probability of a submitter/URM being admitted tends to decrease by about 1 to 2 percentage points in 2018 and 2019, the difference not being statistically significant at the 5% level. The changes in 2020 are larger, and statistically different from the status quo probabilities at the 5% level, but still relatively small in magnitude.

Table 6: Simulated probabilities of admission under full disclosure counterfactual

	2018		2019		2020	
	TO	FD	TO	FD	TO	FD
	Avg. Prob. Admitted	Avg. Prob. Admitted	Avg. Prob. Admitted	Avg. Prob. Admitted	Avg. Prob. Admitted	Avg. Prob. Admitted
	95% CI	95% CI	95% CI	95% CI	95% CI	95% CI
Submitters	0.3683	.3536	0.3357	.3135	0.4415	.4028
	(0.3527, 0.3839)	(.3365, .3695)	(0.3208, 0.3506)	(.2983, .3297)	(0.4258, 0.4571)	(.3873, .4170)
URMs	0.4270	.4160	0.4099	.3975	0.4545	.4289
	(0.4046, 0.4493)	(.3897, .4437)	(0.3890, 0.4309)	(.3732, .4253)	(0.4320, 0.4770)	(.4007, .4533)

4.5.2 Displacement Results

The finding that the *overall* share of admitted students who submitted scores changes little under full disclosure does not necessarily imply that few individuals see their admissions outcomes change. In the extreme, for instance, every single non-submitter who was admitted under test-optionality could be rejected under full disclosure, with an equivalent number of non-submitters rejected under test-optionality newly admitted. Moreover, the distributional effects of the changes in admissions are presumably an important input into many people's welfare evaluations of such a policy.

Table 7 characterizes the "switchers": those students predicted to be admitted under full disclosure who are rejected in the status quo (the "switch ins") as well as the converse group (the "switch outs"). The results suggest an interesting pattern: non-submitters comprise a sizableshare of the "switch outs," suggesting some may not have been admitted under full disclosure. However, a decent portion of the "switch ins" are also non-submitters, dampening the overall effect towards the modest changes documented in Table 5. A similar dynamic appears possibly at play with the share of URM applicants. These results suggest that the marginal applicants between the two admissions regimes are broadly similar.

Table 7: Characterizing the "Switchers"

	2018 Mean (95% CI)	2019 Mean (95% CI)	2020 Mean (95% CI)
Total Share Switching Admit Status	11.91 (11.32, 12.62)	11.89 (11.21, 12.63)	14.85 (14.05, 15.54)
Share Submitting : Switch Ins	70.03 (66.67, 75.34)	75.04 (71.11, 80.00)	73.89 (71.48, 79.81)
Share Submitting: Switch Outs	58.73 (54.78, 64.26)	56.05 (50.64, 60.20)	51.97 (48.22, 57.00)
Average SAT Score: Switch Ins	1253.20 (1241.20, 1262.32)	1267.76 (1258.35, 1279.52)	1277.42 (1269.06, 1289.36)
Average SAT Score: Switch Outs	1208.19 (1195.13, 1223.91)	1204.53 (1190.82, 1215.91)	1194.57 (1183.92, 1205.84)
Average HS GPA: Switch Ins	3.51 (3.49, 3.53)	3.53 (3.51, 3.55)	3.55 (3.53, 3.57)
Average HS GPA: Switch Outs	3.32 (3.30, 3.34)	3.31 (3.28, 3.32)	3.26 (3.24, 3.27)
Share URM: Switch Ins	18.67 (15.94, 22.31)	15.25 (12.56, 18.03)	10.32 (7.61, 12.36)
Share URM: Switch Outs	22.44 (18.06, 26.40)	21.38 (18.22, 24.82)	17.96 (15.32, 21.55)

4.5.3 Assessing the Capacity Constraint

As noted in the introduction, a priori, the general equilibrium effects imposed by the assumption that colleges must maintain the same total number of admits could be an important consideration. Moreover, this would seem a plausible explanation for the muted overall effects of the policy, if large changes in the distribution of admissions ratings force large countervailing changes in the value of the threshold. This section suggests, however, that empirically, these effects are non substantial in my data . The results in Panel A of Table 8 simulate the composition of an admitted class where the threshold does not adapt; in other words, I now predict an applicant to be admitted if and only if their fitted probability from step 4 of the procedure in Proposition 2 exceeds .5. As Panel A shows, the results are overall very similar to those in Table 5, with the confidence intervals largely overlapping. Panel B of Table 8 shows that the overall number of students predicted to be admitted without imposing the capacity constraint is in general very close to the actual number of admits, with the 95% confidence interval on the simulated moment always containing the actual number of admitted students in that year. In a similar vein, Panel C provides the simulated thresholds from the simulations presented in Table 5, showing that they are in general close to .5, with their confidence intervals always overlapping with .5. In addition to assessing the impact of the general equilibrium effects on the composition of admits, these results can be read as a robustness check on the assumption of a binding capacity constraint, suggesting that the assumption of the college being capacity-constrained is not a major determinant of the results.

Table 8: Simulated Full Disclosure Admissions without Binding Capacity Constraint

	2018	2019	2020
Panel A : Composition without capacity constraint	Mean 95% CI	Mean 95% CI	Mean 95% CI
Share Submitting	76.87 (75.59, 78.30)	77.75 (77.19, 78.31)	73.27 (72.51, 74.03)
Avg SAT Score	1305.78 (1300.82, 1310.72)	1306.96 (1300.90, 1313.02)	1294.21 (1292.74, 1295.67)
Avg HS GPA	3.63 (3.62, 3.64)	3.65 (3.64, 3.65)	3.64 (3.63, 3.64)
Share URM	14.13 (12.87, 15.35)	15.55 (14.77, 16.33)	14.05 (12.89, 15.22)
Panel B : # Admitted without capacity constraint			
# Predicted Admitted	5379 (5209, 5528)	5260 (5109, 5419)	5730 (5566, 5973)
# Actually Admitted	5465	5325	5711
Panel C : Threshold with capacity constraint			
Threshold	.4872 (.470, .505)	.4888 (.4714, .5021)	.5053 (.4902, .5215)

4.5.4 Assessing the Impact of Test Scores on Admissions

In addition to the offsetting effects of would-be test-optional students displacing one another, a further reason for the modest effects of a switch to full-disclosure may be that relative to other factors, test scores are simply not very determinate of who is admitted. Table 9 examines this possibility by simulating an extreme case: in each year, using the estimated coefficients from the probit regression on submitters in that year, I fit the predicted probabilities of admission for the non-submitters after changing their scores to be either the lowest among non-submitters in that year or the highest. I then compute, when all test-optional applicants are assessed as if they have the low score, how many of those who are admitted in the status quo are still predicted to be admitted (fitted probability above .5) , and conversely, when all applicants are rated as if they have the high score, how many of those who are rejected in the status quo are still predicted to be rejected (fitted probability below .5). The results show that even when the maximum and minimum scores are extreme in magnitude, a sizable chunk of the admits/rejects have an admissions prediction in line with their status quo outcome.

Table 9: Admissions Under Extreme Scores

	Year		
	2018	2019	2020
Highest score	1460	1450	1600
Share rejects still predicted rejected at highest score	84.29 (82.45, 86.63)	84.71 (82.24, 87.14)	75.42 (69.36, 80.80)
Lowest score	790	490	630
Share admits still predicted admitted at lowest score	81.33 (67.87, 91.82)	47.67 (27.75, 60.61)	66.26 (50.55, 79.79)

A potential explanation for the results in Table 9 can be found in two sets of statistics. The first are those documented in section 3.6.3, showing that the admitted non-submitters tend to fare much better than the rejected non-submitters along other academic dimensions, such as GPA and curriculum score. The second is documented in Table 10 below: after converting test scores to be on the same scale as GPA, the average marginal effect of higher test scores on the probability of admission for the submitters is much lower than that for GPA, suggesting the latter is more important for admissions under either regime. One potential concern is that these results simply suggest that I am underestimating the full impact of test scores on admissions probabilities, with the underlying function being perhaps nonlinear. However, adding in higher-order test-score terms to the regression consistently yields insignificant coefficient estimates and barely increases the model fit.

Table 10: Marginal Effects of Test Score vs GPA on Admission

	Year		
	2018	2019	2020
Test Score	.0655	.106	.079
	(.0344, .0964)	(.0733, .138)	(.0428, .115)
GPA	.665	.630	.595
	(.645, .685)	(.609, .651)	(.571, .619)

Notes: *Test score has been divided by 400 to be on the same scale as GPA.*

5 Testing Models of Test-Optionality

As highlighted in the preceeding section, a key advantage of my approach to simulating a full-disclosure equilibrium is that it bypasses the need to take a stance on how admissions operates in a test-optional environment. Nevertheless, understanding precisely which mechanisms are at play under test-optionality may be useful for evaluating the external validity of my results along at least two dimensions. First, it may help explain how the results would (or would not) generalize across colleges; for instance, if sensitivity to rankings appears to be a major factor in admissions, then the impacts of test-optionality may vary depending on how sensitive a given school is to its rankings, while if "cursed" behavior (i.e. irrationaity on the part of student and/or admissions officers) seems like a major explanation, the comparative sophistication of those actors across schools may be a key source of cross-sectional variation. Second, decomposing the channels through which test-optionality operates could help explain how the results would (or would not) generalize over time, as the policy landscape surrounding test-optional admissions changes. For instance, the impact of rankings may change due to responses on the part of rankings organizations: the US News and World Report historically had designed its formula to partially penalize test-optional schools, but it lessened the penalty during the COVID-19 pandemic due to the higher expected share of students exercising test-optional policies. [9] Even without external policy changes, the outcomes of test-optional admissions could evolve over time if the applicant pool or admissions officers learn from historical performance; for instance, perhaps colleges are initially cursed, neglecting to consider that non-submission indicates lower scores, but eventually examine the underlying scores of admitted non-submitters versus submitters and revise their behavior towards a rational expectations equilibrium.

5.1 A Very Simple Model

I first explore arguably the simplest possible model of test-optionality, and show that it generates "full unraveling": all but the lowest-scoring students submit, even given the choice not to. The point of this model is not to be realistic – the prediction of full unraveling is flatly rejected by the descriptive statistics in section 3.6.1– but to narrow down by contrapositivity the set of assumptions of which at least one must fail for the empirical results to hold.

This model involves overlaying three more assumptions on the baseline setup (Assumption 1–3) from section 4.1. First, I impose further assumptions on the college's rating function.

Assumption 6. *(College preferences and behavior)*

(i) The ratings functions $g(SAT_i, \mathbf{X}_i)$*,* $h(\tilde{SAT}(\mathbf{X}_i), \mathbf{X}_i)$ *are strictly increasing in their respective first arguments, the (actual or imputed) SAT scores.*

(ii) The college cares about scores only as a metric of student caliber, therefore, if a student chooses not to submit, colleges form an imputed test score based off their non-submission choice and their covariates.

(iii) The college is are risk-neutral with respect to student's test score, i.e., they do not weight the application components of non-submitters and submitters any differently.

Part (i) of Assumption 6 may appear innocuous in that colleges would seem to naturally prefer higher-caliber students; it does, however, rule out dynamics such as "yield protection" in which colleges reject high-scoring applicants who they worry will end

[9] In particular, as documented here, "A change for the 2022 edition -- if the combined percentage of the fall 2020 entering class submitting test scores was less than 50 percent of all new entrants, its combined SAT/ACT percentile distribution value used in the rankings was discounted by 15 percent. In previous editions, the threshold was 75 percent of new entrants. The change was made to reflect the growth of test-optional policies through the 2019 calendar year and the fact that the coronavirus impacted the fall 2020 admission process at many schools."

up being admitted to higher-ranked schools and enrolling there instead, lowering the college's yield rate. Part (ii) imposes that colleges care about scores only as a metric of caliber and not, for instance, to maximize rankings that depend on the average scores of submitters (in this sense, assuming part (ii) makes part (i) yet more plausible, since colleges presumably care about maximizing their yield rate only for its effect on rankings). Part (iii) rules out reweighting, i.e., that the imputed scores of non-submitters are weighted less in the admissions function than the observed scores of submitters, with other components weighted more heavily to compensate; this follows under risk-neutrality because the college's imputed scores are conditional expectations of students' scores given their covariates and choice not to submit, which ensures that the imputed score equals the expected score among those applicants, leaving the college indifferent between the two metrics. This lack of reweighting along with the assumption in part (ii) that the college cares only about scores as a metric of caliber implies that $h(\tilde{\text{SAT}}(\mathbf{X}_i), \mathbf{X}_i) = g(\tilde{\text{SAT}}(\mathbf{X}_i), \mathbf{X}_i)$, i.e., the functional form is identical across submitters and non-submitters, with the only difference being whether scores are evaluated at their actual or imputed values.

Next, I make a technical assumption on the distribution of test scores:

Assumption 7. *(Full Support)*

For each possible value $\mathbf{X}_i$ of the covariate vector $\mathbf{X}$, the conditional probability distribution $f(SAT|\mathbf{X}_i)$ has the same support as the unconditional distribution $f(SAT)$, and both are non-trivial (i.e., they are not singletons).

Assumption 7 rules out that there are any classes of students for whom the probability of some scores is precisely 0 or that all students have the same single score; of course, the probability of certain scores may be arbitrarily small. Lastly, I define the informational environment and equilibrium concept.

Assumption 8. *(Equilibrium Concept)*

Students' and the college's behavior as described in Assumptions 1–3 and 6 are common knowledge, and all parties are rational. Therefore, the solution concept is a rational expectations equilibrium with students and the college taking each others' behavior into account.

Under these assumptions, the following holds:

Theorem 1. *The unique caliber-only test-optional equilibrium is for all but the lowest possible score types to submit*

Proof. Consider a given value of the covariate vector $\mathbf{X}_i$. Suppose towards a contradiction that that among the applicants A for whom $\mathbf{X} = \mathbf{X}_i$, there are at least two distinct scores SAT $= j, k \in N(\mathbf{X}_i)$, where $N(\mathbf{X}_i)$ is the set of non-submitters for that type, with $j < k$ without loss of generality . Conditional on $i \in N(\mathbf{X}_i)$, $f(j|i \in N(\mathbf{X}_i), \mathbf{X}_i)$ and $f(k|i \in N(\mathbf{X}_i), \mathbf{X}_i)$ are each strictly positive by assumption 8, which implies that $j < \mathbb{E}[\text{SAT}|\text{SAT} \in N(\mathbf{X}_i), \mathbf{X}_i] < k$. Thus student k prefers to submit, a contradiction. Thus $N(\mathbf{X}_i)$ must be a singleton. Suppose, again towards a contradiction, that $N(\mathbf{X}_i) = \{k\}$ where $k \neq 0$. Then the college would infer a mass point for non-submitters at this score, i.e. $f(k|i \in N(\mathbf{X}_i), \mathbf{X}_i)$. But then all students with scores $j < k$ would find it profitable to deviate to not submitting, so this cannot be an equilibrium. Thus the unique equilibrium non-submission set for any $\mathbf{X}_i$ is the singleton set $\{k\}$ where no students have scores below k: SAT $= 0$. □

I conclude this section with two brief remarks on this result.

Remark 1. If the distribution of scores were continuous, the lowest possible scores would occur with measure zero, making the equilibrium generically one of full submission. In reality, scores are discretely distributed, but the lowest possible score types tend to be extremely rare, such that Theorem 1 would predict the vast majority of students to submit.

Remark 2. It is not necessary for colleges to form the correct conditional distributions $f(\text{SAT}|\mathbf{X})$ over scores or for students to believe they do; so long as the distribution satisfies the full support assumption for each $\mathbf{X}_i$, the proof will go through identically.

5.2 Models of Partial Non-Submission

Theorem 1 is an "unraveling" result in the spirit of Akerlof (1970), except unraveling is the only possible outcome because there are no prices/costs (literal or otherwise), i.e. no market mechanism to potentially abate the adverse selection; in this respect it mirrors the costless signaling results in Grossman (1981). Theorem 1 relies on strong assumptions, and there are many possible ways that relaxing these assumptions could alter the result. In this section, I briefly discuss three modifications to the assumptions underlying the theorem that could plausibly generate the result borne out in the descriptive statistics: some, but not all, students choosing not to submit, i.e. partial non-submission. The point of this section is not to thoroughly model all possible explanations for partial non-submission; in reality, there are likely many other possible explanations, including combinations of the models described below. Rather, the purpose of this section to identify some testable predictions within the data which can allow us to at least narrow down the set of realistic models, ultimately with the aim of strengthening the generalizability of the results in Section 4.

5.2.1 Differential Weights

One simple explanation for partial non-disclosure lies in relaxing the assumption that the college is risk-neutral with respect to student test scores. If the college were instead risk-averse, it may naturally lower the weight placed in the admissions function on the imputed test scores of non-submitters relative to other, correlated factors like GPA. Maintaining the assumption of full rationality, this would in turn incent those with relatively lower scores to not submit for a given value of the covariate vector $\mathbf{X}_i$. The balance between the incentive for low-scorers to shift towards a function that weights their non-score components more highly and the fact that this generates adverse selection that lowers the imputed score of non-submitters could then, in principle, generate an interior equilibrium.

An alternative to microfounding why colleges may reweight the components of the admissions function for non-submitters is to simply take at face value the stated approach of many test-optional schools, which suggest they engage in a particularly extreme form of reweighting that ignores imputed scores entirely. Consider the counsel offered by Yale University to prospective applicants:

> "For applicants without scores, the Admissions Committee places greater weight on other parts of the application, such as high school transcripts, recommendation letters, and essays."[10]

Even more explicitly, 567 test-optional schools signed on to a statement during the COVID-19 pandemic titled "Test-Optional Means Test-Optional" promising that "they will not penalize students for the absence of a standardized test score." [11]

Such statements motivate a basic model for how test-optional applicants might be treated. For simplicity, I model the underlying admissions functions as linear and, without loss, normalize the coefficient on SAT_i to 1. Students are then evaluated according to

[10]This quote can be found here.

[11]The link to this statement and the list of signatories is available here.

$$r(\text{SAT}_i, \mathbf{X}_i, 0) = \begin{cases} \text{SAT}_i + \beta'\mathbf{X}_i & S_i = 1 \\ \theta + \beta^{H'}\mathbf{X}_i & S_i = 0 \end{cases} \tag{5.1}$$

where $\beta^H \gg \beta$ (i.e., β^H is greater in each component). In other words, all test-optional applicants are effectively treated as if they have some shadow score θ, with no attempts by universities to statistically discriminate along their observables, and have their remaining covariates $\mathbf{X}_i$ evaluated at a greater weight. Under Equation 5.1, students will clearly find it optimal to submit according to

$$S_i = \mathbb{I}\{\text{SAT}_i > \theta + (\beta^{H'} - \beta')\mathbf{X}_i\};$$

for proper choice of $\theta, \beta^{H'}$, it is easy to see how this could generate an interior equilibrium; moreover, the equilibrium would satisfy the finding in the descriptive statistics that higher scorers tend to submit.

5.2.2 Cursed Equilibrium

An alternative hypothesis that could generate partial non-submission considers instead the extent to which the involved parties in fact form rational inferences. Intuitively, the assumption that in games of incomplete information agents correctly understand how other parties' actions indicate their hidden information may appear to demand extreme sophistication; applied to our setting, it would require for instance that students correctly hazard that colleges will impute their scores conditional on their choice not to submit, and that in equilibrium, this imputation must result in a score lower than the one they have at any potential threshold, with an analogous mode of higher-order reasoning on the part of the college. Eyster and Rabin (2005) formalize the intuition that such rationality might be psychologically unrealistic by defining a "cursed equilibrium" wherein agents underestimate how other parties' actions (in this context, the choice to submit scores) reveal their types (in this context, the underlying test score). Applied to my setting there are at least two different ways in which a cursed equilibrium could play out: for the college and for students.

Cursed College The college is "fully cursed" in the language of Eyster and Rabin if it neglects to consider the informational signal of the non-submission choice, i.e., if for any non-submitter of type $\mathbf{X}_i$, the college form their imputed score as

$$\tilde{\text{SAT}}(\mathbf{X}_i) = \mathbb{E}[\text{SAT}|\mathbf{X}_i],$$

the unconditional average score of that type. In such an environment, students of type $\mathbf{X}_i$, whether because they themselves neglect the informational content of their signals, or because they are rational but correctly conjecture that colleges are cursed, will submit iff $\tilde{\text{SAT}}(\mathbf{X}_i) > \text{SAT}_i$. Since $\mathbb{E}[\text{SAT}|\mathbf{X}_i] \in (0, 1)$ for all types by Assumption 5, an interior share of each applicant types submits under this model. Furthermore, every applicant's rating is weakly greater under cursed equilibrium than under full disclosure, with strict inequality for at least some applicants.

Cursed Students An alternative possibility is that students behave as if colleges are cursed, but colleges in fact are not and instead form rational inferences. This could occur if either (i) students are aware that colleges would rationally impute scores based off their non-submission choice, but incorrectly conjecture that the college is cursed, or (ii) students themselves neglect to consider that their submission choices carry information regarding their scores, and thus think that the rational behavior of the college is to form expectations unconditional on disclosure. Then, students of type i will submit iff

$$\text{SAT}_i < \mathbb{E}[\text{SAT}|\mathbf{X}_i],$$

and receive imputed scores by the college:

$$\tilde{\text{SAT}}(\mathbf{X}_i) = \mathbb{E}[\text{SAT}|\text{SAT} < \mathbb{E}[\text{SAT}|\mathbf{X}_i]; \mathbf{X}_i].$$

For instance, if for a given $\mathbf{X}_i$, $\text{SAT}_i \sim U[0,1]$, students with scores below .5 will not submit, and their imputed score will be .25 . In general, Assumption 5 implies that an interior share of each applicant types submits, but now the impact on ratings is ambiguous; with some non-submitters students necessarily receiving higher ratings and others necessarily receiving lower ones.

5.2.3 Rankings Maximization

A final potential explanation I consider for the fact that in equilibrium some students choose not to submit – as well as a (cynical) explanation for why colleges may go test-optional in the first place – lies in the role of college rankings systems, such as the US News and World Report's annual list. Universities are generally taken to care greatly about these metrics, as evidenced by recent high-profile accusations that Columbia University falsified its data to boost its ranking. Moreover, the average test score *among matriculating students who submitted scores* comprises roughly 5% of the US News and World Report's ranking system,[12] providing a direct incentive for colleges to care differentially about the (actual) scores of submitters and (imputed) scores of non-submitters.

To formalize this, suppose now that the admissions functions are given by

$$\begin{cases} \alpha \text{SAT}_i + \beta(\text{SAT}_i - \text{SAT}_{\text{max}}) + m(\mathbf{X}_i) & S_i = 1 \\ \alpha \tilde{\text{SAT}}(\mathbf{X}_i) + m(\mathbf{X}_i) & S_i = 0 \end{cases}$$

for some function $m(\cdot)$, i.e., for both submitters and non-submitters, the admissions function is additively separable in (imputed) test score and the other covariates. For submitters, colleges care about both the score as a signal of caliber, with weight α , and the downward effect on rankings with weight β, but for non-submitters, only the caliber effect matters. Note that if $\alpha = 0$, no students but the very highest scorers would submit; I thus impose that both α and β are strictly positive. Since the colleges do care about the true score of the non-submitters as a caliber metric, they form the rational imputation

$$\tilde{\text{SAT}}(\mathbf{X}_i) = \mathbb{E}[\text{SAT}|S_i = 0; \mathbf{X}_i]$$

Optimal student behavior is thus given by $S_i = \mathbb{I}\{\alpha \text{SAT}_i + \beta(\text{SAT}_i - \text{SAT}_{\text{max}}) > \alpha \tilde{\text{SAT}}(\mathbf{X}_i)\}$. This characterization can be sharpened with the following lemma:

Lemma 1. *(Single-Crossing)*

Optimal submission behavior for each $\mathbf{X}_i$ is given by a threshold score $SAT^*(\mathbf{X}_i) \in [0,1]$ such that students of type $\mathbf{X}_i$ submit iff $SAT_i > SAT^*(\mathbf{X}_i)$.

Proof. Recall that submission is determined by a map $S : \mathbf{Z}_i \to [0,1]$ such that $S_i = 1$ iff $\alpha \text{SAT}_i + \beta(\text{SAT}_i - \text{SAT}_{\text{max}}) > \alpha \tilde{\text{SAT}}(\mathbf{X}_i)$. For a given $\tilde{\text{SAT}}(\mathbf{X}_i)$, let $\text{SAT}^*(\mathbf{X}_i) = \inf_{\text{SAT}_i \in [0,1]} \{\text{SAT}_i : \alpha \text{SAT}_i + \beta(\text{SAT}_i - \text{SAT}_{\text{max}}) > \alpha \tilde{\text{SAT}}(\mathbf{X}_i) \geq$

[12] The entire ranking formula is documented here.

$g(\tilde{\text{SAT}}(\mathbf{X}_i), \mathbf{X}_i)\}$. Since the LHS is increasing in SAT , every student of type $\mathbf{X}_i$ with a score above $\text{SAT}^*(\mathbf{X}_i)$ wants to submit, and by definition of infimum, no student of type $\mathbf{X}_i$ with a score below $\text{SAT}^*(\mathbf{X}_i)$ wants to submit. □

If the distribution of scores is continuous, this lemma directly implies that at an interior equilibrium, the equation

$$(\alpha + \beta)\text{SAT}_i^* - \beta\text{SAT}_{\text{max}} = \alpha\mathbb{E}[\text{SAT}|\text{SAT}_i < \text{SAT}^*(\mathbf{X}_i); \mathbf{X}_i]$$

is satisfied. Example 1 shows with a simple parametrization how such an equilibrium could exist.

Example 1. *Assume that $\alpha = 2, \beta = 1$, and SAT $\sim U[0,1]$. Assume that students differ only in their test scores and thus this is the only argument of the admissions function (or equivalently, that all other covariates are independent of the test score distribution) . The rational expectations equilibrium is then given by the threshold score SAT* such that $3SAT^* - 1 = SAT^*$, i.e. $SAT^* = \frac{1}{2}$.*

5.3 Testing Models of Partial Non-Submission

5.3.1 Testing for Full Reweighting

Recall Equation 5.1 derived in Section 5.2.1, the admissions functions that would obtain under many universities' stated behavior of not drawing inferences on underlying test scores and instead fully reweighting towards other variables:

$$r(\text{SAT}_i, \mathbf{X}_i, 0) = \begin{cases} \beta_0\text{SAT}_i + \beta'_{-0}\mathbf{X}_i & S_i = 1 \\ \theta + \beta^{H'}_{-0}\mathbf{X}_i & S_i = 0 \end{cases}$$

Equation 5.1 motivates a simple test for whether this school behaves in the manner many universities describe: namely, running probit regressions on the set of observable covariates for submitters and non-submitters and comparing the coefficients on the non test-score academic variables. If schools are behaving as prescribed by Equation 5.1, θ will simply be soaked up into the constant term for the regression for non-submitters. Table 11 below shows the results of these regressions. Note that the coefficients in Table 11 are the raw output of the probit regressions, not the average marginal/partial effects; as such, they lack any structural interpretation, but are still useful for comparing signs and magnitudes.

The results give reason to be suspect of the validity of the model described by Equation 5.1. While the coefficient on curriculum_score is greater for non-submitters than submitters, the difference is not significant. Conversely, GPA, which we would imagine would be weighed much more heavily in the absence of test scores, actually has a (statistically) significantly higher coefficient among the submitters than the non-submitters. An obvious concern is that we do not have access to the full set of covariates universities observe (i.e., essays, letters of recommendation, extracurriculars), and these omitted variables will also be weighed more heavily for non-submitters if we take universities' statements at face value. This could inflate the variance of the residual in the non-submitters' regression relative to the submitters', attenuating the estimated coefficients towards zero in the canonical manner that can occur in nonlinear models (Yatchew and Griliches (1985)). However, I nonetheless tentatively take the lower coefficient on GPA for the non-submitters as evidence against the model described in 5.1 for two reasons. First, greater weight on the omitted variables of the non-submitters may also bias the estimates on GPA *upwards* if GPA is positively correlated with these omitted variables, counteracting the attenuation bias. Second, even if the change in weight placed on the omitted variables for non-submitters versus submitters is large, we may expect that the change in weight placed on GPA is even

Table 11: Probit Regression of Admissions on Covariates: Submitters vs. Non-Submitters

	Non-Submitters	Submitters	Non-Submitters	Submitters
VARIABLES	admitted	admitted	admitted	admitted
test_score		0.111***		0.127***
		(0.000852, 0.00137)		(0.000997 - 0.00154)
GPA	3.090***	3.338***	3.298***	3.559***
	(2.960, 3.221)	(3.237, 3.439)	(3.156, 3.440)	(3.451, 3.668)
curriculum_score	0.526***	0.488***	0.588***	0.512***
	(0.494, 0.558)	(0.463, 0.513)	(0.553, 0.624)	(0.485, 0.539)
Constant	-12.20***	-14.17***	-15.42***	-16.48***
	(-12.78, -11.63)	(-14.68, -13.66)	(-17.04, -13.80)	(-17.18, -15.78)
Year FE	Yes	Yes	Yes	Yes
Race FE	Yes	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes	Yes
Application Round FE	Yes	Yes	Yes	Yes
Major FE	Yes	Yes	Yes	Yes
State and Country Fixed Effects	No	No	Yes	Yes
Observations	11,293	21,738	11,019	21,485

95% CI in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: *Test scores are divided by 100 for ease of interpreting the coefficients.*

larger; intuitively, it would seem the variable most closely correlated with SAT score and thus the nearest substitute towards which admissions offices would reweight the most.

Note that this test does not rule out that the college engages in reweighting: they may still weight other components of the application more, but also weight the imputed test score, rather than assigning an identical stand-in value θ for all non-submitters. Testing for such reweighting is fundamentally harder as it requires a specification of the imputed score. Nevertheless, these results can refine our understanding by suggesting that at least for this school, if reweighting is the motive for partial non-submission, such reweighting likely still involves students' imputed scores, opening up the door for adverse selection and statistical discrimination.

5.3.2 Testing for Cursedness

The cursed equilibrium described in section 5.2.2 offers its own testable implications in the data. However, at least two comments are in order regarding what kinds of "cursedness" can be tested. First, it is difficult to separately test for students themselves being cursed versus students correctly assuming that colleges are cursed, although the first explanation would seem the more parsimonious. Secondly, it is challenging to test directly for cursedness with respect to submitting scores, i.e., whether or not students rationally internalize the degree to which their submission choices inform colleges of the distribution of their true score. Rather, this section will test for cursedness on other dimensions, i.e., whether students internalize that other covariates than their submission decision (like GPA) also likely inform the conditional distribution of their scores, and thus the imputed scores colleges would rationally form.

The test for this form of cursedness is simple: I analyze the determinants of SAT scores and how these comove with students' submission choices. The idea is as follows: in a rankings-maximization model, if higher values of a covariate are associated with higher (lower) expected scores, then, conditional on students' test scores, they should be less (more) likely to submit as that covariate increases. Formally, let $\mathbf{X}_i = x_i$ (i.e. there is only one non-test score covariate) and let $\mathbb{E}[\text{SAT}|\text{SAT} \leq \text{SAT}^*(x_i); x_i]$ be increasing in x_i (i.e, the imputed score for non-submitters is increasing in x_i).[13] Then, the following holds:

$$(\alpha + \beta)\text{SAT}_i^* - \beta\text{SAT}_{\text{max}} \leq \mathbb{E}[\text{SAT}|\text{SAT} \leq \text{SAT}^*(x_i); x_i] \implies$$

$$(\alpha + \beta)\text{SAT}_i^* - \beta\text{SAT}_{\text{max}} \leq \mathbb{E}[\text{SAT}|\text{SAT} \leq \text{SAT}^*(x_i'); x_i'], \forall x_i' > x_i$$

i.e. any score such that types x_i' would submit is a score where types x_i would submit, but not vice versa. In reality, with multiple covariates, this implies that, integrating over the other covariates to yield the total share of students with a given covariate equal to x_i who submit, this share will increase in x_i. One would expect a similar result to hold in any model of partial nonsubmission in which students and colleges are rational and colleges care about non-submitters' true scores, such that they are assessed according to an imputed score. Conversely, in a cursed equilibrium, if students are cursed regarding the information content of a covariate x_i, the coefficient on a regression on x_i from a probit regression of submission choices on test scores and $\mathbf{X}_i$ should be insignificant: because students neglect that variable's informational content, the only manner in which x_i affects submission choices is through its possible correlation with test scores, which is controlled for.

Table 12 below shows which variables we should expect students to be sensitive to in their submission choices across a range of specifications. Perhaps unsurprisingly, academic variables like GPA and curriculum score are associated with large increases in expected test scores. Note that in this regression, I use only those students with observed scores rather than imputing scores for some non-submitters, since the focus is no longer on simulating the general equilibrium effects of the policy. Additionally, note that is it not necessary for the covariates to be exogenous, since the importance of these correlations is for estimating the process by which universities may statistically discriminate, not the causal impact on test scores. Instead, the requisite condition is that the regressors are uncorrelated with application components that are observable to the admissions office but not to the econometrician. While it is possible that variables like essay strength and extracurriculars are correlated with test scores, one would expect this correlation to be weaker than with the actual academic variables.

[13] A sufficient condition for this property to hold for the conditional expectation is if the conditional distribution of scores given x_i' hazard-rate dominates that for x_i whenever $x_i' > x_i$.

Table 12: Test Scores by Covariates

VARIABLES	(1) test_score	(2) test_score	(3) test_score	(4) test_score
GPA	96.78***	86.59***	88.22***	85.30***
	(92.19, 101.4)	(82.60, 90.59)	(84.35, 92.08)	(81.47, 89.12)
curriculum_score	34.91***	34.89***	35.36***	34.10***
	(33.48, 36.35)	(33.61, 36.18)	(34.15, 36.58)	(32.90, 35.31)
Constant	682.2***	700.1***	665.0***	700.3***
	(668.7, 695.8)	(683.2, 717.0)	(646.7, 683.2)	(602.3, 798.2)
Year FE	Yes	Yes	Yes	Yes
Application Round FE	No	Yes	Yes	Yes
Race FE	No	No	Yes	Yes
Gender FE	No	No	Yes	Yes
Major FE	No	No	No	Yes
State and Country FEs	No	No	No	Yes

Robust 95% CIs in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: *Test scores are divided by 100 for ease of interpreting the coefficients.*

Table 13 below shows how students' submission choices vary with these covariates after conditioning for test score. The results suggest that students are not cursed along the academic dimensions, with consistently negative signs for GPA and curriculum score. The results on the demographic variables are more mixed. As shown in Tables 17 and 18 in Appendix A (the detailed versions of Tables 12 and 13 respectively), many of the race and gender dummies are highly significant in the regression of test scores on covariates, yet do not have significant coefficients in the regression of submission choices on test scores and those covariates, although the signs are generally opposite between the two regressions, as the theory would predict.

The results in Table 13 could be biased if submission choices are influenced by unobserved factors, however note that (i) this would simply shift the non-cursedness onto these variables, with students not submitting because of better essays, extracurriculars and the like, and (ii) as previously mentioned, one would expect these factors to be far less informative for imputing scores than the academic variables. An alternative hypothesis that is more difficult to rule out is that differential submission owe to different behavioral patterns (i.e. different rates of "cursedness") which are correlated with students' covariates.

Table 13: Submission Choices by Covariates and Test Score

VARIABLES	(1) submitted_scores	(2) submitted_scores	(3) submitted_scores	(4) submitted_scores
test_score	0.665***	0.696***	0.699***	0.714***
	(0.00637, 0.00693)	(0.00671, 0.00721)	(0.00672, 0.00727)	(0.00685, 0.00743)
GPA	-0.514***	-0.483***	-0.479***	-0.516***
	(-0.588, -0.439)	(-0.557, -0.409)	(-0.555, -0.403)	(-0.595, -0.437)
curriculum_score		-0.0743***	-0.0770***	-0.0702***
		(-0.0993, -0.0493)	(-0.102, -0.0515)	(-0.0968, -0.0436)
Constant	-4.919***	-5.528***	-5.618***	-6.339***
	(-5.198, -4.641)	(-5.854, -5.201)	(-5.989, -5.248)	(-7.072, -5.606)
Year FE	Yes	Yes	Yes	Yes
Application Round FE	No	Yes	No	No
Race FE	No	No	Yes	Yes
Gender FE	No	No	Yes	Yes
Major FE	No	No	No	Yes
State and Country FEs	No	No	No	Yes

95% CIs in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: *Test scores are divided by 100 for ease of interpreting the coefficients.*

5.3.3 Testing for Rankings Maximization

In this section, I analyze the last model discussed in Section 5.2, in which concern for rankings on the part of the college results in partial non-submission. I first summarize the existing approach of Conlin, Dickert-Conlin, and Chapman (2013) and show how it can be formalized as the result of the model in Section 5.2.3, before turning to an issue with their approach and promoting a robustness check that extends their methodology, finding broadly similar results.

Summary and Replication of the Conlin, Dickert-Conlin, and Chapman (2013) Approach Conlin, Dickert-Conlin, and Chapman (2013), who have micro data from a (presumably different) college assume that a student's latent admissions rating y_i^* is given by

$$\begin{cases} y_i^* = \delta_s + \delta_a \text{SAT}_i + \mathbf{X}_i \delta + \varepsilon_i & S_i = 1 \\ y_i^* = \delta_{ns} + \delta_p \tilde{\text{SAT}}(\mathbf{X}_i) + \mathbf{X}_i \delta + \varepsilon_i & S_i = 0 \end{cases}$$

with admissions $\varepsilon_i \sim \mathcal{N}(0,1)$ and admissions given by $\text{admit}_i = \mathbb{I}\{y_i^* \geq 0\}$. They then assume that $\tilde{\text{SAT}}(\mathbf{X}_i) = \text{SAT}_i, \forall i$, i.e., colleges perfectly impute the missing scores of non-submitters. Under this assumption, they regress

$$y_i = \delta_s \times S_i + \delta_{ns} \times (1 - S_i) + \delta_a \text{SAT}_i \times S_i + \delta_p \text{SAT}_i \times (1 - S_i) + \mathbf{X}_i \delta + \varepsilon_i$$

and find that $\delta_a > \delta_p, \delta_{ns} > \delta_s$, which they take as a sign of colleges' sensitivity to rankings. Table 14 below replicates their approach and finds a similar pattern in my data: the fixed effect for being a submitter is negative, while the coefficient on test scores is significantly greater for non-submitters.

Table 14: Testing for Ranking Sensitivity: Perfect Imputation

	(1)
VARIABLES	admitted
$\mathbb{I}$\{submitted_scores=1\}	-1.318***
	(-2.051, -0.584)
$\mathbb{I}$\{submitted_scores=0\}×test_score	0.0335
	(-0.0271, 0.0941)
$\mathbb{I}$\{submitted_scores=1\}×test_score	0.115***
	(0.0896, 0.141)
GPA	3.312***
	(3.217, 3.407)
CurriculumScore	0.486***
	(0.462 - 0.509)
Constant	-12.68***
	(-13.46, -11.90)
Year FE	Yes
Race FE	Yes
Gender FE	Yes
Application Round FE	Yes
Major FE	Yes

95% CIs in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Test scores are divided by 100 for ease of interpreting the coefficients.

Microfounding and Extending the Conlin, Dickert-Conlin, and Chapman (2013) Approach There are at least two reasons why one might take issue with this approach. The first is that, while it makes intuitive sense that rankings-sensitivity may lead to the pattern documented above, one may wonder if it in fact follows directly from a model where colleges care about rankings. Here, I show it to follow from the model in section 5.2.3. In particular, we can rewrite the model from that section as

$$y_i^* = \beta_0 S_i + \beta_0 (1 - S_i) + \beta_1 \text{SAT}_i(S_i) + \beta_1 \tilde{\text{SAT}}(\mathbf{X}_i)(1 - S_i) + \gamma(\text{SAT}_i - \text{SAT}_{\text{max}}) \times S_i + \mathbf{X}_i \beta + \varepsilon_i \tag{5.2}$$

A priori, the coefficient on terms 3 and 4 is identical because colleges care equally about scores as a metric of student caliber regardless of whether they are observed or inferred. The intercept term β_0 is also a priori identical for the two classes of students (independent of scores and other covariates, there is no reason colleges prefer one class of students to the others). The fifth term reflects the difference between a submitting student's score and the highest possible submitters; this is the rankings maximization term that reflects the negative effect on rankings imposed by students with imperfect scores. We assume that $\beta_1, \gamma > 0$. Now, regrouping like terms , we arrive at

$$y_i^* = (\beta_0 - \gamma \text{SAT}_{\text{max}})S_i + \beta_0(1 - S_i) + (\beta_1 + \gamma)\text{SAT}_i(S_i) + \beta_1 \tilde{\text{SAT}}(\mathbf{X}_i)(1 - S_i) + \mathbf{X}_i\beta + \varepsilon_i$$

Or equivalently

$$y_i^* = \delta_s S_i + \delta_{ns}(1 - S_i) + \delta_a \text{SAT}_i(S_i) + \delta_p \tilde{\text{SAT}}(\mathbf{X}_i)(1 - S_i) + \mathbf{X}_i\delta + \varepsilon_i$$

where

$$\delta_s = \beta_0 - \gamma \text{SAT}_{\text{max}}, \delta_{ns} = \beta_0;$$
$$\delta_p = \beta_1, \delta_a = \beta_1 + \gamma;$$
$$\delta = \beta$$

A second, and arguably greater, concern with their approach lies in the assumption that colleges can perfectly impute the underlying score of students. This assumption would seem suspect both empirically and theoretically. Empirically, there may be considerable variation in scores even accounting for the rich set of covariates at colleges' disposal; my regressions of test scores on observables tend to have maximum R^2 values less than .5. Theoretically, if colleges could in fact perfectly estimate the scores of non-submitters, everyone but the highest score types would choose not to submit: the imputed score and actual score would obtain the same value in the ratings function, with all scores less than SAT_{max} incurring a penalty from the rankings component; the rankings model relies on the trade-off between the adverse selection of not submitting and the penalty for imperfect scores for submitting to generate an interior equilibrium.

I propose an admittedly heuristic extension to the Conlin, Dickert-Conlin, and Chapman (2013) approach by providing a few different specifications of how admissions officers might impute the scores of non-submitters; in particular, I provide a minimal imputation function wherein they use few variables to form their conditional expectations, a medium one wherein they use some but not all, and a maximal one in which they use almost all. For each of these specifications, I regress students' scores on the specified predictor variables Γ_i , each interacted with the indicator for whether students submitted or not, and then fit the predicted values for the non submitters. To make the results as directly comparable to Table 14 as possible, I only impute scores for non-submitters who in fact have test scores, since those are the observations for which the regression in Table 14 is feasible. Unfortunately, there is a key constraint on the ability to model the imputation function as fully as would be rational for the college to do: if admissions officers conditioned on all possible information, they would form separate estimates for every possible interaction of all the non test-score covariates; however, in my data, there are enough covariates that this would result in a huge number of observations uniquely identified by such an interaction. As such, going from relatively fewer to more terms in the parametrization of the imputation function is an admittedly imperfect proxy for college behavior, assuming it is fully rational.

Table 15 summarizes the results from the three different models of the imputation function, with the variables Γ_i used for the respective imputation listed below the coefficient estimates in each column. The results suggest some movement in the coefficient when going from the minimal to medium imputation functions, but little from going to medium to maximal. Moreover, the basic pattern of Table 14 holds: a significant negative coefficient on the fixed effect for submitters, and a significantly greater coefficient for test scores of the submitters than for the (imputed) test score of the non-submitters. Again, while this is a necessarily imperfect approach to proxying for colleges imputations, it provides tentative support for the hypothesis that rankings-sensitivity determines some of the difference in behavior under test-optionality.

Table 15: Testing for Ranking Sensitivity: Imperfect Imputation

	(1)	(2)	(3)
VARIABLES	admitted	admitted	admitted
$\mathbb{I}$\{submitted_scores=1\}	-3.437***	-2.462***	-2.451***
	(-5.247, -1.628)	(-3.795, -1.128)	(-3.693, -1.209)
$\mathbb{I}$\{submitted_scores=0\}×test_score	-0.160*	-0.0722	-0.0714
	(-0.321, 0.00208)	(-0.191, 0.0472)	(-0.182, 0.0395)
$\mathbb{I}$\{submitted_scores=1\}×test_score	0.111***	0.111***	0.111***
	(0.0852, 0.136)	(0.0855, 0.137)	(0.0852, 0.137)
GPA	3.334***	3.324***	3.324***
	(3.238, 3.431)	(3.228, 3.419)	(3.228, 3.419)
curriculum_score	0.492***	0.490***	0.490***
	(0.468, 0.516)	(0.466, 0.514)	(0.466, 0.514)
Constant	-10.65***	-11.57***	-11.57***
	(-12.40, -8.896)	(-12.89, -10.24)	(-12.82, -10.32)
Year FE	Yes	Yes	Yes
Race FE	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes
Application Round FE	Yes	Yes	Yes
Major FE	Yes	Yes	Yes
Imputation Variables			
GPA	Yes	Yes	Yes
Curriculum_Score	Yes	Yes	Yes
Year	Yes	Yes	Yes
Race	No	Yes	Yes
Gender	No	Yes	Yes
Race×Gender	No	No	Yes
Application Round	No	No	Yes
Observations	24,330	24,330	24,330

95% CIs in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: *Test scores are divided by 100 for ease of interpreting the coefficients. The variable test_score is equal to a student's actual test score if they submitted, and otherwise is equal to their imputed score following the imputation scheme specificed by the bottom panel of this table.*

6 Conclusion

This paper examines the equilibrium effects of test-optional admissions policies on the composition of admitted students. Using proprietary data from a test-optional college, I simulate the effect of a switch to full-disclosure, finding that little would change on aggregate in terms of the share of URM students and the share of students who choose not to submit in the test-optional status quo. My findings suggest two explanations for these small effects: (i) "canceling out" whereby demographics which are well-represented among the students admitted under test-optionality but not full disclosure are also well-represented among the students admitted under full-disclosure but not test-optionality, and (ii) low sensitivity of the admissions function to test scores in general, making other dimensions along which students vary more determinate of their admissions outcome under either regime. One may wish to go further and ask precisely how universities behave under test-optionality to understand the effects of the simulated policy change; while I show that modeling test-optional admissions is non-trivial (and unnecessary for the counterfactual exercise), I test some admittedly strong models to narrow down what we can say about test-optional behavior. Firstly, I note that we can reject a very basic model of test-optional admissions that would predict full unraveling from simple descriptive statistics documenting a sizable share of non-submitting applicants. I then consider some stylized models that can account for such a phenomenon of partial non-submission. My results suggest that (i) contrary to the stated policy of many universities, the college for which I had dada did not simply reweight admissions entirely towards other components for the non-submitters, (ii) applicants to this college are not "cursed" in terms of neglecting the informational content of their high school academic performance, and (iii) the college may be sensitive to rankings that rely on the scores of submitting matriculants only, potentially driving some of the behavior under test-optionality.

The strongest assumption throughout the paper – and the one likely most fruitful to relax – is that students' choices of whether to initially apply to a college is invariant to test-optional policy. Examining departures from this assumption could help assess the robustness of the results in this paper. However, such an analysis would be both theoretically and empirically challenging. Theoretically, introducing a first stage of the model in which applicants choose whether or not to decide would create a complex dynamic problem, since application choices are presumably affected by perceived admissions probabilities, and admissions probabilities in turn depend on the equilibrium distribution of applicants. Empirically, one could imagine two ways of assessing the impact of application choices, neither of which is straightforward: the first would be if a researcher had access to micro data from a college before and after a switch to or from test-optionality, and then specified a structural model of application choices depending on variables like the cost of sending scores, the overall cost of applying, and predicted probabilities of admission conditional on applying. In principle, one could then examine which of the applicants who applied in one admissions regime are predicted to still apply in the other, and then analyze their predicted admissions probability in that regime conditional on applying. However, relevant variables such as effective application costs for students may not be observable in the data, since students qualify for fee waivers heterogeneously, and specifying such a model introduces many of the aforementioned theoretical complications. A second approach would be a differences-in-differences style analysis that examines applicants in test-optional versus full disclosure school(s) before and after the policy change relative to similar school(s) that did not switch their policies. As discussed however, differences-in-differences regarding test-optional admissions that rely on publicly available aggregate data tend to be underpowered and require parallel trends assumption that may be suspect in this setting. One could alternatively run such an analysis using micro data from multiple schools over the same periods, only some of which switched their testing policies; given the difficulty of acquiring any microdata sets on college applicants however, this would be logistically challenging.

Beyond the assumption of unchanged application choices, my paper suggests a number of potential other extensions on both theoretical and empirical dimensions. Theoretically, the models of college behavior under test-optionality explored in Section 5 could be extended in a number of ways. First, one could consider whether the implications of allowing multiple forces considered in this paper to interact, e.g., if colleges are rankings-sensitive and either students or colleges are to some degree cursed. Second, one could consider other models that predict partial non-submission; for instance, models where students are sensitive to the costs of sending scores. Third, one could examine the implications of further dimensions of heterogeneity among applicants beyond their admissions-relevant covariates. For instance, perhaps applicants are heterogeneously cursed, with some neglecting the informational content of their non-submission choice (or their non-test-score observables) more than others; this could introduce potentially novel strategic channels: for example, if some students are fully cursed while the remaining applicants are entirely rational and aware of their peers' cursedness, these rational students may nevertheless find it strategic to not submit, since the action of the cursed students effectively "shades" the imputed scores of the non-submitters towards the unconditional average.

Empirically, the framework laid out in Section 4 could be applied by any researcher with microdata on all applicants to a given college. While my data are remarkably comprehensive, there are some further desirable features that could make such an analysis more straightforward: for instance, if one had access to scores for all non-submitters (say, via merging with data from the College Board or ACT Corporation), they could bypass the imputation step in my procedure. Furthermore, if one had data where a college provides a complete ranking of applicants (such that the researcher effectively observes the latent variable in the binary choice model), one could avoid the necessary complications regarding the variance and normality assumptions that the probit model (or any alternative model such as logit) necessitates. Lastly, if one had data on not just admissions decisions but further outcome variables on enrolled students such as college GPAs or postgraduate earnings, they could further examine the implications of test-optional admissions. Such an analysis would perhaps be another relevant input into policy makers' normative evaluations of test-optionality.

Lastly, note that this paper has compared test-optional policies to full-disclosure admissions under the logic that the former is rapidly increasing its foothold as an alternative to the latter, the historical status quo. However, an interesting question for future research is to examine other potential policies regarding standardized tests in admissions. For instance, "test-blind" policies such as those recently adopted by the University of California system (in which colleges do not consider standardized tests for any applicants) would heighten any effects of reweighting towards other metrics in the absence of scores, but would eliminate any potential adverse selection effects present in test-optionality. Alternatively, expansion of fee waivers would mirror the potential cost-reducing effects of test-optionality. In general, few or sometimes no institutions may have introduced such alternative policies of interest, but given the rapidly changing landscape of college admissions, we may expect new admissions policies with their own uncertain effects to proliferate in coming years.

References

Akerlof, George. 1970. “The Market for ”Lemons”: Quality Uncertainty and the Market Mechanism”. *Quarterly Journal of Economics* 84 (3): 488–500.

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

Bennett, Christopher T. 2022. “Untested Admissions: Examining Changes in Application Behaviors and Student Demographics Under Test-Optional Policies”. *American Educational Research Journal* 59 (1): 180–216.

Blair, Peter Q, and Kent Smetters. 2021. *Why Donât Elite Colleges Expand Supply?* Working Paper, Working Paper Series 29309. National Bureau of Economic Research.

Bleemer, Zachary. 2021. *Top Percent Policies and the Return to Postsecondary Selectivity*. Working Paper.

Callaway, Brantly, and Pedro H.C. Sant’Anna. 2021. “Difference-in-Differences with multiple time periods”. Themed Issue: Treatment Effect 1, *Journal of Econometrics* 225 (2): 200–230.

Chade, Hector, Gregory Lewis, and Lones Smith. 2014. “Student Portfolios and the College Admissions Problem”. *Review of Economic Studies* 81 (3): 971–1002.

Conlin, Michael, and Stacy Dickert-Conlin. 2017. “Inference by college admission departments”. *Journal of Economic Behavior & Organization* 141 (C): 14–28.

Conlin, Michael, Stacy Dickert-Conlin, and Gabrielle Chapman. 2013. “Voluntary disclosure and the strategic behavior of colleges”. *Journal of Economic Behavior Organization* 96:48–64.

Dynarski, Susan, et al. 2021. “Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students”. *American Economic Review* 111 (6): 1721–56.

Eyster, Erik, and Matthew Rabin. 2005. “Cursed Equilibrium”. *Econometrica* 73 (5): 1623–1672.

Goodman-Bacon, Andrew. 2021. “Difference-in-differences with variation in treatment timing”. Themed Issue: Treatment Effect 1, *Journal of Econometrics* 225 (2): 254–277.

Grossman, Sanford J. 1981. “The Informational Role of Warranties and Private Disclosure about Product Quality”. *The Journal of Law Economics* 24 (3): 461–483.

Hoxby, Caroline, and Christopher Avery. 2012. *The Missing ”One-Offs”: The Hidden Supply of High-Achieving, Low Income Students*. NBER Working Papers 18586. National Bureau of Economic Research, Inc.

Kapor, Adam. 2020. *Distributional Effects of Race-Blind Affirmative Action*. Working Paper.

Pallais, Amanda. 2015. “Small Differences that Matter: Mistakes in Applying to College”. *Journal of Labor Economics* 33 (2): 493–520.

Robinson, Michael, and James Monks. 2005. “Making SAT scores optional in selective college admissions: a case study”. *Economics of Education Review* 24 (4): 393–405.

Roth, Jonathan. 2022. “Pre-test with Caution: Event-Study Estimates after Testing for Parallel Trends ”. *American Economic Review: Insights* Forthcoming.

Saboe, Matt, and Sabrina Terrizzi. 2019. "SAT optional policies: Do they influence graduate quality, selectivity or diversity?" *Economics Letters* 174 (C): 13–17.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". Themed Issue: Treatment Effect 1, *Journal of Econometrics* 225 (2): 175–199.

Sweitzer, Kyle, A. Emiko Blalock, and Dhruv Sharma. 2018. "The effect of going test-optional on diversity and admissions: A propensity score matching analysis", 288–308.

Yatchew, Adonis, and Zvi Griliches. 1985. "Specification Error in Probit Models". *The Review of Economics and Statistics* 67 (1): 134–139.

Appendix A

Additional Tables and Graphs

Table 16: Composition of Submitters and Non-Submitters by Race/Ethnicity and Gender

Panel A: Composition By Race

	Submitted Scores		
	N	Y	Total
	Frequency	Frequency	Frequency
Race/Ethnicity	(Percent)	Percent)	(Percent)
Asian	295	553	848
	(2.58)	(2.53)	(2.55)
Black or African American	1,085	591	1,676
	(9.49)	(2.71)	(5.04)
Hispanic of any race	2,023	2,152	4,175
	(17.69)	(9.86)	(12.56)
Native American or Alaskan Native	12	16	28
	(0.10)	(0.07)	(0.08)
Native Hawaiian or Pacific Islander	6	8	14
	(0.05)	(0.04)	(0.04)
Nonresident Alien	530	630	1,160
	(4.63)	(2.89)	(3.49)
Not Provided	151	302	453
	(1.32)	(1.38)	(1.36)
Race/Ethnicity Unknown	151	334	485
	(1.32)	(1.53)	(1.46)
Two or more races	311	538	849
	(2.72)	(2.47)	(2.55)
White	6,874	16,691	23,565
	(60.10)	(76.51)	(70.87)
Total	11,438	21,815	33,253

Panel A: Composition By Gender

	Submitted Scores		
	N	Y	Total
Gender	Frequency	Frequency	Frequency
	(Percent)	Percent)	(Percent)
F	7,201	11,367	18,568
	(62.96)	(52.11)	(55.84)
M	4,237	10,448	14,685
	(37.04)	(47.89)	(44.16)
Total	11,438	21,815	33,253

Table 17: Test Scores Given Covariates: Detailed Breakdown

	(1)	(2)
VARIABLES	test_score	test_score
Academic Variables		
Curriculum_Score	35.36***	34.10***
	(34.15, 36.58)	(32.90, 35.31)
GPA	88.22***	85.30***
	(84.35, 92.08)	(81.47, 89.12)
Race		
Black or African American	-100.1***	-94.87***
	(-110.7, -89.49)	(-105.1, -84.63)
Hispanic of any race	-54.84***	-48.55***
	(-63.71, -45.97)	(-57.21, -39.90)
Native American or Alaskan Native	-21.32	-26.43
	(-70.31, 27.67)	(-73.64, 20.79)
Native Hawaiian or Pacific Islander	19.48	10.56
	(-47.42, 86.39)	(-53.84, 74.95)
Nonresident Alien	-14.13**	-43.11***
	(-25.20, -3.065)	(-58.57, -27.65)
Not Provided	29.07***	21.81***
	(15.25, 42.88)	(8.447, 35.18)
Race/Ethnicity Unknown	22.66***	17.63***
	(9.540, 35.77)	(4.917, 30.35)
Two or more races	11.67**	11.83**
	(0.293, 23.05)	(0.818, 22.84)
White	12.79***	9.318**
	(4.650, 20.93)	(1.413, 17.22)
Gender		
Male	43.00***	37.50***
	(40.37, 45.63)	(34.70, 40.31)
Constant	665.0***	700.3***
	(646.7, 683.2)	(602.3, 798.2)
Year FE	Yes	Yes
Application Round FE	Yes	Yes
Race FE	Yes	Yes
Gender FE	Yes	Yes
Major FE	No	Yes
State and Country FEs	No	Yes

95% CI in parentheses

*** p<0.01, ** p<0.05, * p<0.1

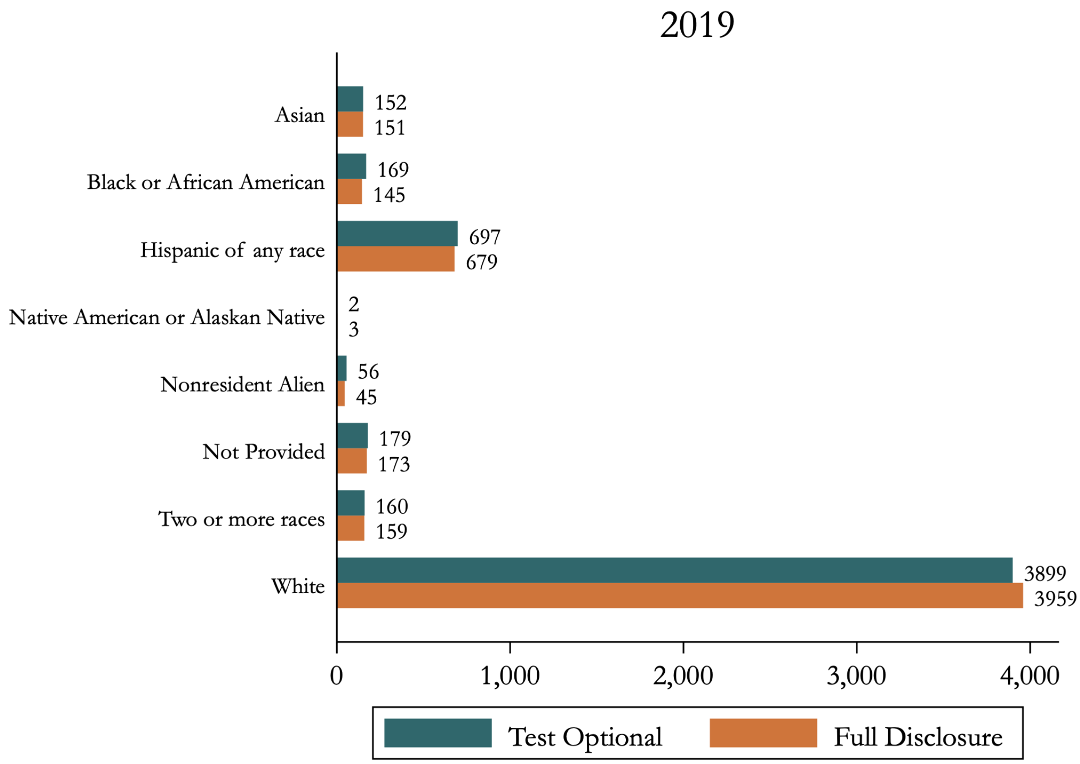
Table 18: Probability of Submission Given Test Scores and Other Covariates: Detailed Breakdown

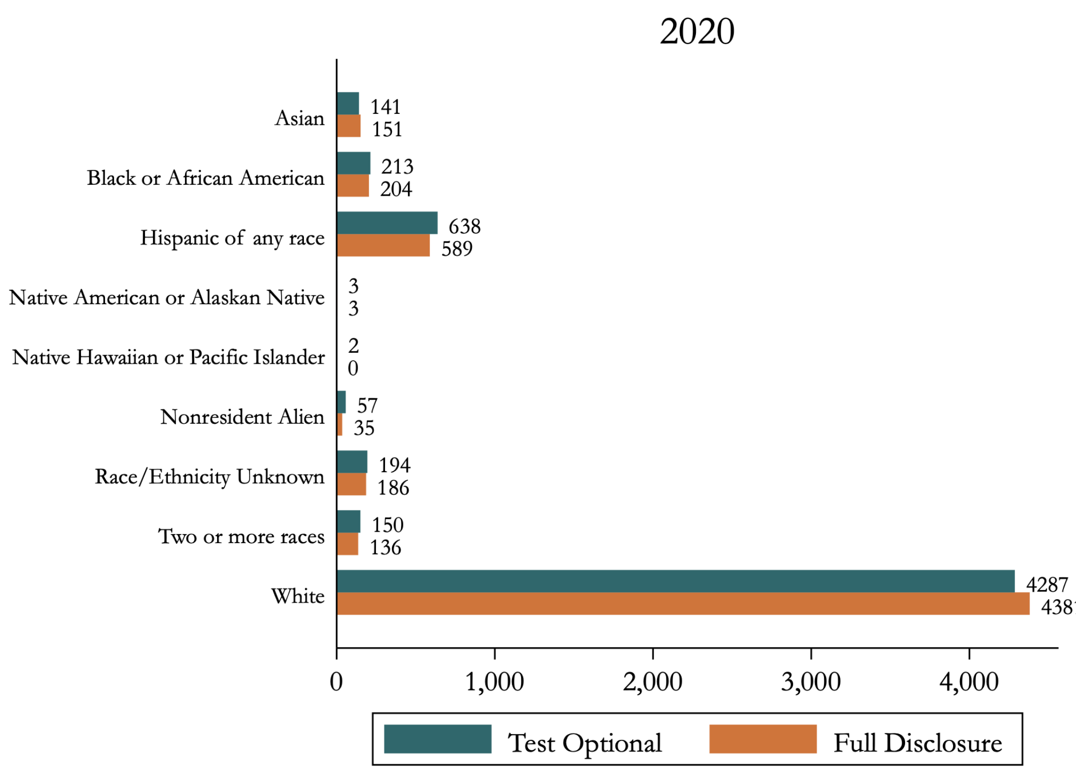
	(1)	(2)
VARIABLES	submitted_scores	submitted_scores
Academic Variables		
test_score	0.00699***	0.00714***
	(0.00672, 0.00727)	(0.00685, 0.00743)
Curriculum_Score	-0.0770***	-0.0702***
	(-0.102, -0.0515)	(-0.0968, -0.0436)
PCGPA	-0.479***	-0.516***
	(-0.555, -0.403)	(-0.595, -0.437)
Race		
Black or African American	0.0326	0.0184
	(-0.154, 0.219)	(-0.172, 0.209)
Hispanic of any race	0.0866	0.0638
	(-0.0801, 0.253)	(-0.108, 0.236)
Native American or Alaskan Native	1.003	1.080*
	(-0.251, 2.257)	(-0.162, 2.322)
Native Hawaiian or Pacific Islander	0.0124	0.0124
	(-1.353, 1.378)	(-1.417, 1.442)
Nonresident Alien	0.191*	0.394**
	(-0.0278, 0.409)	(0.0826, 0.706)
Not Provided	0.0160	0.0353
	(-0.302, 0.334)	(-0.292, 0.363)
Race/Ethnicity Unknown	0.0707	0.0490
	(-0.191, 0.332)	(-0.217, 0.315)
Two or more races	0.000800	0.00687
	(-0.221, 0.223)	(-0.219, 0.233)
White	0.0261	0.0220
	(-0.130, 0.182)	(-0.138, 0.182)
Gender		
Male	0.0373	0.0179
	(-0.0168, 0.0914)	(-0.0423, 0.0781)
Constant	-5.618***	-6.339***
	(-5.989, -5.248)	(-7.072, -5.606)
Year FE	Yes	Yes
Application Round FE	No	No
Race FE	Yes	Yes
Gender FE	Yes	Yes
Major FE	No	Yes
State and Country FEs	No	Yes
Observations	24,336	24,030

95% CI in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 3: Number of Admits By Racial/Ethnic Classifications: Test Optional versus Full Disclosure, 2019 and 2020 Results





Appendix B

Robustness Checks for Full-Disclosure Simulation

Alternative Specification of the Admissions Function Given my focus on the impacts of a switch to full disclosure on the racial and ethnic composition of the admitted class, one concern may be if the admissions function in fact weights test scores differently for different demographic groups. While any fixed difference in how scores are assessed for one racial/ethnic group versus another will be soaked up via the fixed effect for that group, perhaps not just the intercept but the slope of the admissions function with respect to test scores changes via racial/ethnic classification. Table 19 below assesses robustness to this possibility by repeating the procedure in Proposition 2 but with the admissions function in Step 3 now regressin on an interaction term between test score and IPEDS racial/ethnic classification in place of simply regressing on the test score, with all the other regressors identical. The results are highly similar to the baseline results in Table 5. I also show the threshold for these new simulations to stress that the results do not appear to owe to differential general equilibrium effects relative to the baseline.

Table 19: Simulated Composition of Admits under Full Disclosure with Test Score / Race Interaction

	2018 Mean 95% CI	2019 Mean 95% CI	2020 Mean 95% CI
Share Submitting	76.60 (75.27, 77.95)	77.64 (76.34, 78.95)	73.37 (72.04, 74.70)
Avg SAT Score	1304.54 (1299.73, 1308.98)	1305.08 (1300.39, 1309.76)	1294.07 (1289.97, 1298.17)
Avg HS GPA	3.62 (3.62, 3.63)	3.65 (3.64, 3.65)	3.64 (3.63, 3.65)
Share URM	14.36 (13.23, 15.74)	15.51 (14.31, 16.71)	13.33 (12.14, 14.53)
Threshold	.4863 (.4677, .5034)	.4907 (.4744, .5068)	.5030 (.4873, .5188)

Alternative Imputation of Missing Scores Another concern one may have with the procedure in Proposition 2 regards the validity of the imputation procedure for missing scores. Perhaps those non-submitters with scores visible in the data differ from those without scores not just observably but unobservably, biasing the predictions obtained by steps 1 and 2 of my procedure. Tables 20 and 21 show the results of a heuristic exercise to assess robustness to my baseline results, wherein I instead impute the missing scores "optimistically" and "pessimistically" respectively. The optimistic imputation fits the missing scores using the coefficients of a regression of test scores on the observables of all students with visible scores, including students who submitted them. Perhaps unsurprisingly, this results in much higher estimates of the missing scores, increasing the average scores of admitted students, but does not change the other results meaningfully relative to the baseline in Table 5.

Table 20: Simulated Composition of Admits under Full Disclosure with Optimistic Imputation

	2018 Mean 95% CI	2019 Mean 95% CI	2020 Mean 95% CI
Share Submitting	76.78 (75.49, 78.06)	76.58 (75.45, 77.80)	72.52 (71.33, 73.80)
Avg SAT Score	1303.57 (1298.77, 1308.65)	1324.74 (1321.54, 1328.39)	1309.96 (1289.97, 1298.17)
Avg HS GPA	3.62 (3.62, 3.63)	3.65 (3.64, 3.65)	3.64 (3.63, 3.65)
Share URM	14.13 (12.95, 15.34)	15.78 (14.56, 16.89)	14.04 (12.64, 15.05)
Threshold	.4862 (.4692, .5045)	.4907 (.5074, .5232)	.5143 (.5023, .5308)

Conversely, in Table 21, I impute the missing scores using the coefficients of a regression of test scores on the observables of all non-submitters with visible scores who were rejected in the status quo. Again, the results change little from the baseline.

Table 21: Simulated Composition of Admits under Full Disclosure with Pessimistic Imputation

	2018 Mean 95% CI	2019 Mean 95% CI	2020 Mean 95% CI
Share Submitting	75.85 (74.44, 76.90)	77.91 (76.67, 79.34)	73.25 (71.99, 74.63)
Avg SAT Score	1324.67 (1321.39, 1327.63)	1301.74 (1269.74, 1306.57)	1291.89 (1287.38, 1296.73)
Avg HS GPA	3.62 (3.62, 3.63)	3.64 (3.64, 3.65)	3.64 (3.63, 3.65)
Share URM	14.36 (13.29, 15.67)	15.56 (14.43, 16.77)	13.98 (12.64, 14.99
Threshold	.4998 (.4847, .5174)	.4876 (.4714, .5049)	.5051 (.4908, .5243)

Appendix C

Additional Proofs

Proof of Proposition 2

Under the econometric assumptions I impose, plim$\hat{\beta} = \beta$, so I show that the procedure is valid for the true coefficients. Since by Propositon 1 there exists a unique threshold π' that prevails under full-disclosure equilibrium, it follows that a student is predicted to be admitted if $\Pr(\beta' \mathbf{z}_i \geq \pi') \geq \frac{1}{2}$. For the students with visible scores (whether test-optional or not) , this can be written as $\Pr(\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i + \varepsilon_i \geq \pi') \geq \frac{1}{2}$, where β_{-0} refers to the coefficient vector excluding the coefficient β_0 . Since $\varepsilon_i \sim \mathcal{N}(0,1)$, we can equivalently write this as

$$\Pr(\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i - \pi' \geq \varepsilon_i) = \Phi(\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i - \pi')$$

If $\Phi(\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i - \pi') \geq \frac{1}{2}$, then $\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i \geq \pi'$ after taking the inverse normal CDF of both sides by monotonicity. Taking the normal CDF of both sides of this rearranged inequality then yields that equivalently,

$$\Phi(\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i) \geq \Phi(\pi')$$

.

Meanwhile, for the students with imputed scores , we have

$$\Pr(\beta_0 \text{SAT}_i + \beta'_{-0}\mathbf{X}_i + \varepsilon_i \geq \pi') = \Pr\left(\beta_0(\gamma'\mathbf{X}_i + \nu_i) + \beta'_{-0}\mathbf{X}_i - \pi') \geq -\varepsilon_i\right)$$

$$= \Pr\left(\beta_0 \gamma'\mathbf{X}_i + \beta'_{-0}\mathbf{X}_i - \pi') \geq -\varepsilon_i - \beta_0\nu_i\right)$$

$$= \Pr\left(\beta_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i - \pi' \geq \varepsilon_i + \beta_0\nu_i\right)$$

by symmetry of the two error terms. Note furthermor that $\varepsilon_i, \nu_i | \mathbf{X}_i$ are normally distributed each with mean zero and respective variancse 1 and $\sigma_{\nu|\mathbf{X}_i}$. Letting their correlation be $\rho_{\mathbf{X}_i}$, the standard result on the sum of correlated normals yields that $\varepsilon_i + \hat{\beta}_0\nu_i \sim \mathcal{N}(0, 1 + \beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i})$. Thus

$$\Pr\left(\beta_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i - \pi' \geq \varepsilon_i + \beta_0\nu_i\right) = \Pr\left(\frac{\beta_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i - \pi'}{\sqrt{1 + \beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i}}} \geq \frac{\varepsilon_i + \nu_i}{\sqrt{1 + \beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i}}}\right)$$

$$= \Phi\left(\frac{\hat{\beta}_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i - \pi'}{\sqrt{1 + \beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i}}}\right)$$

Note, though, that

$$\Phi\left(\frac{\beta_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i - \pi'}{\sqrt{1 + \beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i}}}\right) \geq \frac{1}{2} \iff \frac{\beta_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i - \pi'}{\sqrt{1 + \beta_0^2\sigma^2_{\nu|\mathbf{X}_i} + 2\rho_{\mathbf{X}_i}\beta_0\sigma_{\nu|\mathbf{X}_i}}} \geq 0$$

Since the denominator is strictly positive, this is furthermore equivalent to

$$\hat{\beta}_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i \geq \pi' \iff \Phi(\hat{\beta}_0 \text{S}\hat{\text{A}}\text{T}_i + \beta'_{-0}\mathbf{X}_i) \geq \Phi(\pi')$$

by monotonicity. Thus, it suffices to search for the $\Phi(\pi') \in [0,1]$ such that the number of applicants with fitted probabilities above that threshold equals the number admitted in the status quo, whether that applicant has imputed or actual scores. And since no threshold between any two fitted probabilities will yield a different number of predicted admitted students, it is sufficient to simply find the Nth highest predicted probability, where N is the number of students admitted in the status quo.

Appendix D

Data Construction Details for Graph of Share Test-Optional Schools Over Time

Initial Construction and Cleaning

I use data from the IPEDS from 2006–2019 on all schools that are "Degree-granting, primarily baccalaureate or above" (i.e., 4-year colleges and universities; there are 2,045 such institutions). I filter out all schools that are for-profit, that use open admissions policies, and that have more than 2 missing years in the time period.

Coding Test Optionality

Currently, there are four non-NA options in the IPEDS survey in response to the question of whether an institution is test-optional:

- 1: Required
- 2: Recommended
- 3: Neither Required nor recommended
- 5: Considered but not required

(I consider "Recommended" to mean effectively required for the purposes of the graph). Option 5 most closely matches to what is mean by test-optional and schools are explicitly advised in an FAQ on the IPEDS website to answer with this response if they are test-optional. However, in the panel I have, the earliest year any school lists option 5 is 2016. I take this to mean before that then option 5 did not exist in the survey (the other three options go back to the beginning of the panel). However, in the absence of option 5 existing there are two sets of schools that could plausibly list option 3: test-optional schools (schools that will consider your score if submitted but do not require it) and test-blind schools (schools that ignore scores entirely). Some schools (in particular, 171 universities) list option 3 up to 2016 and then start listing 5 or list 3 up to a few years after 2016, the laggards likely owing to the novelty of the 5 option and being used to filling out the form with 3. It seems likely that these schools were test-optional as long as they had been listing 3, because it is as far as I know unprecedented for a school to go from test-blind to test-optional. I then drop schools from this subsample of schools who list 5 at some point that (i) were test-optional for the entire period or (ii) reverted from test-optional to not test optional.

The ambiguity is with schools that continuously list 3 all the way up to 2019, of which there are 98. For many of these schools a voluntary information form known as the Common Data Set (CDS) is also available. In the CDS there is both a "consider if submitted" and a "do not consider" option for SAT/ACT going all the way back to 2006, and schools seem to recognize this difference. I manually inspect the 98 schools in question and determine their test optionality based on the CDS data if available, and if not, I use university press releases / websites as archived on the Wayback Machine. The results of this exercise are available here. It results in 65 schools being dropped because they either (i) had unreliable data (contradictions across sources), (ii) were test-optional for the entire period , (iii) reverted from test-optional to not test optional or (iv) were in fact test-blind. Of the remaining 33 schools, 31 are test-optional the first time they list "3" in IPEDS and remain test-optional for the entire rest of the sample, while 2 are test-optional the second time they list "3;" this likely owes to confusion the schools had over whether to use the admissions policy for the coming cycle or previous cycle in answering the IPEDS survey.