NBER WORKING PAPER SERIES

VOLUNTARY REPORT OF STANDARDIZED TEST SCORES: AN EXPERIMENTAL STUDY

Marty Haoyuan Chen
Ginger Zhe Jin

Working Paper 33660
http://www.nber.org/papers/w33660

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2025

We are thankful to the University of Maryland College of Behavioral and Social Sciences for the generous funding used for subject payments. We would like to thank Nolan Pope, Emel Filiz-Ozbay, Erkut Ozbay, Melissa Kearney, David Deming, Caroline Hoxby, Zachary Bleemer, and seminar and conference participants at 2024 NBER Summer Institute on the Economics of Education, the 2025 Disclosure, Information Sharing, and Secrecy (DISS) Workshop, University of Maryland, University of Buffalo, and Colgate University for helpful comments and discussions. This project was approved by the IRB at the University of Maryland. All rights reserved, all errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Voluntary Report of Standardized Test Scores: An Experimental Study
Marty Haoyuan Chen and Ginger Zhe Jin
NBER Working Paper No. 33660
April 2025
JEL No. D61, D63, D8, I23, I24

ABSTRACT

The past few years have seen a shift in many universities' admission policies from test-required to either test-optional or test-blind. This paper uses laboratory experiments to examine students' reporting behavior given their application package and the school's interpretation of non-reported standardized test scores. We find that voluntary disclosure is incomplete and selective, supporting the incentives of both partial unraveling and reverse unraveling. Subjects exhibit some ability to learn about the hidden school interpretation, though their learning is imperfect. Using a structural model of student reporting behavior, we simulate the potential tradeoff between academic preparedness and diversity in a school's admission cohort. We find that, if students have perfect information about the school's interpretation of non-reporting, test-blind is the worst and test-required is the best in both dimensions, while test-optional lies between the two extremes. When students do not have perfect information, some test-optional policies can generate more diversity than test-required, because some students with better observable attributes may underestimate the penalty on their non-reporting. This allows the school to admit more students that have worse observable attributes but report. We test the results' robustness to a variety of extensions.

Marty Haoyuan Chen
University of Maryland
hchen915@umd.edu

Ginger Zhe Jin
University of Maryland
Department of Economics
College Park, MD 20742-7211
and NBER
ginger@umd.edu

1 Introduction

In the past few years, many universities have dropped their SAT and ACT requirements, switching to a test-optional or test-blind admission procedure. While most schools initially did so to accommodate applicants during the worst of the COVID-19 pandemic, the shift away from score requirements has become popular even as threats from the pandemic recede. Education experts — including economists, other social scientists, policy makers, and practitioners — debate whether dropping the SAT and ACT requirements contributes to better admission outcomes, especially in light of the education inequalities across applicants of different family and economic backgrounds.

This paper focuses on one important but understudied factor in this debate: when students have the option to disclose or hide their standardized test scores, this decision is likely strategic. Schools' admission office may or may not take such strategic behavior into account, but students' strategic decision would depend on their belief of how the admission office would interpret non-reporting, which in turn affects the final admission outcomes.

Existing evidence already points to the potential importance of selective reporting. According to Freeman, Magouirk and Kajikawa (2021), the percent of students reporting SAT score in college application has declined sharply from 73% during the 2019-2020 season to 40% in 2020-2021 while the percent of common app member colleges that did not require test scores rose from roughly one-third to 89%. This is opposite to the classical disclosure theory (Milgrom, 1981; Grossman, 1981), which predicts that a rational receiver of the disclosure signal should assume all non-reported students have the worst test scores and therefore all students except for those with the worst scores should report. Clearly, the reality is far from the unraveling equilibrium.[1] Given such selection, it is not surprising that most colleges ranked top 100 by the US News & World Report have seen their distribution of SAT scores of the admitted class improve in the past few years, because this distribution is conditional on the admitted students that had reported SAT scores to the college. [2]

One may argue that strategic reporting is of little importance because standardized test score is only one of many student attributes that admission officers may consider in a

[1]Take the University of Texas Austin as an example. Its press release on March 11, 2024 states that 42% of its freshman applications for Fall 2024 reported their standardized scores and the median SAT score of these reporting students is much higher than that of those who did not report (1420 versus 1160). Source: https://news.utexas.edu/2024/03/11/ut-austin-reinstates-standardized-test-scores-in-admissions/.

[2]Figures A.1 and A.2 show the trends in SAT and ACT scores in the last few years. Among applicants who submitted their SAT/ACT scores and got admitted to the top-100 schools ranked by U.S. News, there was an increasing trend in average test scores in the past few years.

college application. Admission officers may be able to guess the non-reported test scores because these observable factors are correlated with test scores. However, such correlation is often imperfect. Growing evidence suggests that standardized test scores such as SAT and ACT are highly predictive of students' academic performance in college as well as career earnings even after controlling for other observable student attributes (Bettinger, Evans and Pope, 2013; Sanchez and Comeaux, 2020; Chetty, Deming and Friedman, 2023; Cascio et al., 2024). Some of these findings — or the logic behind them — may have driven schools like MIT, Dartmouth, Yale, Brown, and the University of Texas at Austin to return to the test-required policy. However, the vast majority of colleges are still test-optional, some schools (such as University of California) have gone to the extreme of test-blind (namely not accepting any score reporting even if the student volunteers to disclose it). These ongoing developments motivate us to study the incentive and consequences of voluntary reporting of standardized test scores.

The key questions we ask in this paper are three-fold: First, how do students choose to report or not report their standardized test score when they observe their own score in private but recognize that their other application materials (GPA, family economic status, activities, etc.) are observable to the college? Second, how would students' strategic reporting behavior affect the college's admission outcome in terms of academic preparedness and diversity of admitted students? Third, how should a college choose among test-required, test-blind and test-optional policies if it appreciates both academic preparedness and diversity of the admitted class?

We answer these questions in a clean lab setting where a number of human subjects (applicants) play a simple college admission game for 50 rounds. In each round, each subject receives a private endowment A and a public endowment B. The endowments are randomly drawn with a positive correlation between A and B. Each student's B is automatically observable to the college but students choose whether to report A to the college. Upon student choice of reporting or non-reporting, the college (simulated by computer) admits half of the applicants that it believes to have the best total endowment ($A + B$).

A simplified college admission game highlights two strategic incentives in student reporting. First, students with a higher score are more likely to report, and this incentive is stronger when the college interprets non-reported score more harshly. If the school interprets non-reporting as the worst score possible, everyone would have an incentive to report except for those of the worst score, leading to the classical unraveling equilibrium. However, if the school interprets non-reporting strictly above the worst score (for an ideological reason for example), only those that have a score above the school's interpretation of non-reporting would have an incentive to disclose. We refer to this threshold-based incentive as "partial unraveling."

While the partial unraveling incentive focuses on the applicant's own score, the second incentive of strategic reporting also depends on the other application materials that the college observes on the applicant. This is modeled as the public endowment in our model and experiment. To the extent that the school believes in a positive correlation between the public and private endowments of the student, it will infer a non-reporting student's private endowment based on her public endowment. This conditional interpretation introduces a "reverse unraveling" incentive because students with better public endowment would expect more favorable interpretation of non-reporting by the school, which in turn discourages her from reporting.

Not only does the reverse unraveling incentive reduce information available to a test-optional school and therefore force it to rely more on the public endowment in admission decision, it but also affects the distribution of the admitted class. More specifically, students with high public endowment but low private endowment (e.g. high-SES-low-achieving students) can better hide behind their high public endowment under test-optional than test-required, at the expense of low public endowment high private endowment (low-SES-high-achieving) students. This leads to less diversity as measured by the standard deviation of the public endowment of admitted students and less academic preparedness as measured by the average private endowment of admitted students.

Test-blind is even worse than test-optional in both dimensions, because it deprives any opportunity of low-SES-high-achieving students standing out via voluntary report of test score, and maximize the favorable mask on high-SES-low-achieving students. The lack of information on private endowment implies that a test-blind school has to rely on public endowment only to predict each applicant's total endowment. Given the positive correlation between private and public endowments, it ends up admitting students with the highest public endowment. This reduces diversity and academic preparedness of the admitted class, relative to all test-optional and test-required policies.

In short, our illustrative model suggests that the perceived tradeoff between academic preparedness and diversity does not exist if students are all rational and have perfect information on the school's interpretation of non-reporting. Test-required would Pareto dominate test-optional in both dimensions, which further dominates test-blind. Of course, in reality, students may not be fully rational and may not have perfect information on the school's interpretation of non-reporting. To mimic the real world, we informed our lab subjects that A and B are randomly drawn but positively correlated, we allowed subjects to learn their own admission outcome in each round, we also told subjects the 25th and 75th percentiles of the private endowment of the students that reported their private endowment and got admitted in the last round (akin the US News Report in reality). By design, the subjects do not know how the school interprets non-reporting exactly but they can learn about it round by round through the observed admission outcomes and admission distributions. Within this framework, we test two test-optional policies (T1 and T2), with T2 being more lenient than T1 in the school's interpretation of non-reporting.

Despite the imperfect information, we observe strong evidence in support of partial unraveling and reverse unraveling. In particular, when the school interprets non-reporting as having a private endowment equal to half of the 25th percentile of the disclosed private endowment of admitted students in the last round (our T1 treatment), the average reporting rates increase monotonically from 21.7% to 93.5% if the subject's private endowment A goes from the lowest (1) to the highest (5). Most of these reporting rates decrease if the school interprets non-reporting more leniently as having a private endowment equal to the average of the 25th and 75th percentiles of the reported private endowment among students admitted in the last round (our T2 treatment). These patterns suggest that higher-score students are more likely to report, especially when the school interprets non-reporting harshly, confirming the partial unraveling incentive.

Furthermore, within the same treatment, if we compare students with exactly the same private endowment (say $A = 3$) but different public endowment, those with the lowest public endowment ($B = 1$) are most eager to report (89.6% in T1 and 81.9% in T2), but the vast majority of those with the highest public endowment ($B = 5$) are reluctant to report (only 21.4% report in T1 and 11.9% report in T2). The sharp drop of reporting rate by public endowment confirms the reverse unraveling incentive.

Because the reverse unraveling incentive counters the partial unraveling incentive and the two endowments are positively correlated, the unconditional probability of reporting (across all values of private endowment in the same treatment) is still higher for those with the highest public endowment. This is consistent with the real world statistics reported by the University of Texas Austin – the SAT reporting rate is higher for students in the top 6% of their high school class than all applicants in total (49% vs. 42%), although the majority in both groups choose non-reporting. This also suggests that a simple comparison of reporting rate by public endowment alone can be misleading because even if the reverse unraveling incentive cancels out the partial unraveling incentive on average, their seemingly comparability masks important distributional changes driven by strategic reporting.

To enhance our understanding of subjects' belief of the school's admission policy, we estimate a logit model of subject disclosure decision. We use both a Heuristic model and a

Bayesian updating model to capture subject learning at any point in the experiment. The average distance between the actual reporting rates and the model predicted reporting rates is roughly 1.2-1.5 percentage points.

In addition to the partial and reverse unraveling incentives, we find evidence that subjects learn round by round about the hidden admission policy, but in the meantime, there is a small probability that subjects follow a naive rule of thumb of not reporting until their private endowment is strictly above the population average regardless of the admission history they can observe in the lab. These imperfections in rationality and information set do not overturn the insights from the model with perfect information. When we use the empirically estimated parameters to simulate reporting decision and admitted outcomes across 16 test-optional policies, we find test-required performs better in both the average preparedness and diversity of the admitted students than 11 of the 16 test-optional policies, and test-blind is always the worst in both dimensions. The few test-optional policies that are not dominated by test-required demonstrate more diversity but lower academic preparedness than test-required. This tradeoff occurs because our subjects do not possess perfect information of the school policy. In particular, when the school imposes a harsh penalty on non-reporting but some students with better observable attributes underestimate this penalty, the school can admit more students that have worse observable attributes but report. When we simulate under perfect information, these two test-optional policies are also dominated by test-required in both dimensions.

Finally, using our structural model, we extend the stylized college application problem to examine the tradeoff between academic preparedness and diversity when we allow for applicant-group admission quota, resource constraints in test-taking, or an additional dimension of student application profile that signals the student's academic ability besides the standardized test score. In most cases, our previous findings remain robust: test-required is the best for academic merit and demonstrates little sacrifice in diversity than most versions of test-optional, while test-blind is the worst in both dimensions. However, when schools impose a strict interpretation of non-reporting, test-optional policies can introduce a meaningful tradeoff between average preparedness and diversity as compared to test-required, especially if standardized test score is a noisy measure of academic merit. As in our baseline model, this tradeoff disappears if students have perfect information about the school's interpretation of non-reporting, and standardized test score (albeit noisy) is sufficiently informative of the student's true merit. When standardized test score is very noisy and less informative than the school's other signal about student merit, our simulation suggests some room for test optional policies to increase diversity of the admitted class, usually at the expense of average academic preparedness.

Our findings have far-reaching policy implications. For higher education, the findings suggest that, as long as the market can learn from cohort to cohort and standardized test score is sufficiently informative about student's academic ability, it is difficult to justify test-optional or test-blind as a policy better than test-required, as far as the goal of admission is to improve the merit and diversity of the admitted class and regardless of how the school values these two admission outcomes relatively. One may argue that college admission is a once-a-lifetime decision for most students, and thus imperfect information is more realistic than perfect information; in that case, the potential gain of diversity under test-optional with imperfect information may be worth pursuing even if it may reduce the average academic preparedness of the admitted class. This argument is subject to two caveats. First, while perfect information is hard to achieve in reality, we do observe meaningful learning in the lab. In reality, key market players — including parents, high schools, and private counseling services — all have strong incentives to learn and disseminate the information, especially when the decision could be life-changing. In such an environment, the potential gain of diversity from test-optional is hardly sustainable in the long run. Second, the test-optional policies that could bring some gain in diversity tend to put harsh penalty on non-reporting (although not as harsh as the unraveling theory describes). This harshness is at odds with the messages from many test-optional schools that try to convince students that non-reporting will not put students at a disadvantage. If schools' admission practice is consistent with these messages, our results suggest that there may not be much gain of diversity in the first place.

Related literature. Our study contributes to two strands of the literature. First, we add to the literature on the role of standardized test in college admissions. While test-optional policies were largely concentrated in selective liberal arts colleges, the last decade has witnessed the expansion of test-optional adoption to an abundance of schools of various types. This expansion was further accelerated by the COVID-19 pandemic. Early papers have empirically studied the effects of test-optional policies using pre-pandemic data (Belasco, Rosinger and Hearn, 2015; Saboe and Terrizzi, 2019; Bennett, 2022). They find limited effect of test-optional policies on increasing the application volume and the diversity of enrolled students, including the proportion of Pell Grant recipients enrolled. Test-optional policies did result in higher reported SAT scores, which could boost the ranking of schools (Dynarski et al., 2023). A recent empirical study analyzes applicants to 50 major U.S. colleges for entry in Fall 2021, and find strategic disclosure of test scores among these applicants (McManus, Howell and Hurwitz, 2023). Consistent with our findings, they find applicants withheld low scores and disclosed high scores, and that their disclosure choices are dependent on their other academic characteristics, colleges' selectivity and testing policy statements. They do not find large differences in test disclosure strategies by applicants' race and socioeconomic status.

In addition to these empirical studies, there is also a theoretical literature on test-optional college admissions. Borghesan (2022) develop an equilibrium model that allows applicants to endogenously determine their test-taking and school-application decisions, and colleges to adjust admissions thresholds to maximize their objectives. The model predicts reduced student quality at elite schools and negligible increase in college attendance for low-income students under a test-blind policy. While reduced standardized-test weight in college admission seems unappealing from previous studies, Dessein, Frankel and Kartik (2023) argue that social pressure could justify test-optional policies. They propose a model in which college disagrees with the society on the desired composition of admitted students, and show that a test-optional policy could help college reduce the "disagreement cost" with society. Related to our discussion of the tradeoff between academic preparedness and diversity, Liang, Lu and Mu (2021) study the tradeoff between accuracy and fairness in a broader context. They show that excluding test scores is welfare-reducing as long as group identity (e.g., race) is a permissible input in admission decisions, while it might be preferred with an affirmative action ban. Finally, (Garg, Li and Monachou, 2020) develop a model where the school can design their admissions procedure and choose the information that it requires the applicants to submit. They find that eliminating standardized test scores may improve welfare in the presence of the effect of access barriers on the applicant pool size. They further provide a threshold characterization regarding when removing a feature improves both academic merit and diversity.

Our study is different from the above test-optional literature in that we use lab experiments to exclude endogeneity concerns, while focusing on applicants' strategic reporting behavior. This allows us to elicit subject beliefs, construct a model of applicant reporting decision, and simulate reporting behavior and admission outcome for any given counterfactual school policies. Nevertheless, our model does not capture the cost of test preparation and other barriers to college applications.

Beyond college admissions, standardized exams are used for screening in many other contexts. Most related to our work is Moreira and Pérez (2022), who study the impact of the 1883 Pendleton Act. They find that the introduction of competitive exams increased the representation of individuals with high education but limited connections, and reduced the share of lower-socioeconomic status federal employees selected.

Second, our study is related to the literature on voluntary disclosure of verifiable information. The classical unraveling results suggest that the same outcome from mandated disclosure can be achieved if the (voluntarily) disclosed information is verifiable and the related costs are close to zero (Milgrom, 1981; Grossman, 1981). In practice, however, voluntary disclosure is far from complete in many industries (Dranove and Jin, 2010; Jin, Luca and Martin, 2021; Feltovich, Harbaugh and To, 2002; Eyster and Rabin, 2005; Board, 2009; Hirshleifer and Teoh, 2003). The voluntary disclosure of standardized test is no different: less than half of college applicants submitted SAT or ACT scores in the year of 2022-2023.[3] Unraveling does not arrive in our model because school's belief about non-disclosed test scores does not degenerate to the worst possible score and is dependent on the applicant's non-test characteristics. When the quality is multi-dimensional and the voluntarily disclosed element is correlated with other elements, the receiver (school) does not necessarily interpret non-reporting as the worst. This disincentivizes the disclosure of non-favorable information.

This paper proceeds as follows. Section 2 defines the college application problem and uses an illustrative model to compare the admission outcomes of test-required, test-blind and test-optional policies. Section 3 outlines the experiment design. Section 4 discusses the results from the experiments. Section 5 presents a structural model of subject reporting decision and Section 6 presents a welfare analysis. Section 7 concludes.

2 The College Application Problem

In this section, we first describe a simplified college application problem and discuss the predicted admission outcomes under test-required, test-blind, and test-optional policies. This illustrative model aims to highlight students' strategic choice of score reporting and the important role that the school's interpretation of non-reporting plays in this process. Then we outline the college application problem in our experiments, which allows subjects to have imperfect information on the school's interpretation of non-reporting.

2.1 A Simplified Problem

Setting. Consider a single-college application problem, with N student applicants. Each student's application profile has two components: a private endowment A (standardized test scores) and a public endowment B (e.g., high school GPA, letters of recommendation, extracurricular activities). The student observes her own private and public endowments and chooses to either report or not report the private endowment. All students

[3]Source: https://s3.us-west-2.amazonaws.com/ca.research.publish/Deadline+Updates/DeadlineUpdate-_030223.pdf.

understand that private and public endowments are positively correlated in the applicant population, and everyone's public endowment is automatically observable to the college. Without loss of generality, we can rewrite A as:

$$A = \alpha B + e \tag{1}$$

where $\alpha > 0$ and e is independent of B. In words, e represents the "new" information in a student's A that cannot be inferred from her public endowment B. The student observes her own A and B and thus e, but the college cannot observe her A or e unless she reports A. After each student makes the reporting decision, the college takes a guess on each student's private endowment, and admits $N/2$ students based on each student's perceived total endowment T. For simplicity, we assume B and e conform to a uniform distribution between 0 and 1 independently.

School. The school admits students based on $T = \bar{A} + B$, where $\bar{A}$ is the school's belief of a student's private endowment based on the student's reporting decision (R). If A was reported ($R = 1$), $\bar{A}$ is equal to the true A. If A was not reported ($R = 0$), $\bar{A}$ would be given by the function $g(B)$, which takes the form:

$$g(B) = \alpha B + c \tag{2}$$

where c is a constant and α is the same as the α in Equation 1 because the positive correlation between A and B is assumed to be public knowledge. One can also interpret α as the probability under which the school interprets non-reported private endowment as perfectly identified by the student's public endowment; otherwise, the school interprets non-reported private endowment as a constant ($c/(1-\alpha)$). Either way, higher α denotes stronger correlation between the two endowments, and therefore less new information contained in private endowment conditional on public endowment. Given α, higher c implies that the school would interpret the non-reported A more leniently. We can summarize the school's expectation of A as:

$$\bar{A} = \begin{cases} A & \text{if } R = 1 \\ \alpha B + c & \text{if } R = 0. \end{cases} \tag{3}$$

Student. If students have perfect information on the school-belief function $g(B)$ and they act fully rational to maximize the admission probability, they should report if their private endowment is higher than school's expectation of non-reported private endowment. In other words, they follow the reporting rule:

$$R = \mathbb{1}\Big\{A > \alpha B + c\Big\} = \mathbb{1}\Big\{e > c\Big\} \tag{4}$$

where the student reports if her private endowment A is above the threshold $\alpha B + c$. Put it another way, conditional on the student's public endowment B, she would only report if the new information contained in her private endowment is above the school's interpretation of this new information upon non-reporting ($e > c$).

Two incentives are worth highlighting. First, students with higher A are more likely to report A. This selection has been well documented in the classical unraveling literature: higher quality firms are more motivated to disclose their product quality to the public because their true quality exceeds consumer interpretation of non-disclosed quality. Conditional on α and B, the threshold of reporting ($\alpha B + c$) increases with c, which suggests that a more lenient interpretation of non-reporting would motivate more students to hide the score. When the school has the harshest interpretation ($c = 0$), every one discloses because $e > 0$ everywhere, leading to the classical unraveling equilibrium. But as long as $c > 0$, students with $e < c$ would choose non-reporting. We refer to this incentive as "partial unraveling." More lenient interpretation of non-reporting leads to less partial unraveling.

The second incentive of strategic reporting is less obvious: since public and private endowments are positively correlated ($\alpha > 0$), students with higher public endowment (higher B) face a higher reporting threshold ($\alpha B + c$). Put it another way, if a student earns a relatively high but not full score in the standardized test, she is more reluctant to report the score if she comes from a high-income family, enrolls in a good high school, has high GPAs, etc. This happens because the school would interpret her non-reported score more favorably based on her high public endowment, a logic we refer to as "reverse unraveling."[4]

In our model, partial unraveling and reverse unraveling cancel out each other, because we assume α is public knowledge and e is independent of B by definition. This leads to an overall reporting rule of $e > c$ regardless of B. If the school's belief of α (or the students' understanding of the school's belief of α) is different from the actual α, the two incentives

[4]Note that this is different from the counter-signaling effect shown in Feltovich, Harbaugh and To (2002); Bederson et al. (2018). Counter-signaling is more complicated than reverse unraveling, because to ensure counter-signaling as a subgame perfect equilibrium one needs the disclosed signal to be coarser than the true quality and the presence of another exogenous but noisy signal. Otherwise, the equilibrium with rational expectation boils down to classical unraveling. Here we shy away from multiple and coarse signals but allow the school's interpretation of non-reporting to be exogenous and non-rational so that the interpretation may not coincide with the average private endowment of those who choose non-reporting.

may not completely cancel out each other.

Admission outcome. Figure 1 shows the composition of the admission cohort under different admission policies. In each of the six panels, the square represents a uniform distribution of e and B in the applicant population. Panels (a), (b) and (c) present the admission outcome under test-required, test-blind, and test-optional, respectively. Panel (d) highlights the difference between test-required and test-blind; Panel (e) highlights the difference between test-required and test-optional; and Panel (f) highlights the difference between two test-optional policies of different leniency.

Let us first consider the test-required policy. Because we assume $T = A + B = (1+\alpha)B+e$, the indifferent curve that represents a particular value of T is a downward sloping straight line with slope $-(1+\alpha)$. To maximize total endowment of admitted students, the school would find an indifferent curve that represents the population median of total endowment and admit every student with a total endowment above it. This corresponds to the blue shaded area in Panel (a). More specifically, the school would reject anyone with $B < \frac{1}{2} - \frac{1}{2(1+\alpha)}$ regardless of their test score (referred to as "straight reject"), accept anyone with $B > \frac{1}{2} + \frac{1}{2(1+\alpha)}$ (referred to as "straight accept"), and trade off between B and e for any students in between (referred to as "tradeoff group"). A higher α implies less new information in test score, which makes the T-indifferent curve steeper and therefore expands the straight reject and straight accept groups. As a result, test score matters for fewer students in the tradeoff group.

Following the same logic, when a school adopts a test-blind policy, it can only admit students based on expected total endowment, which can be written as $E(T|B) = E((1+\alpha)B + e|B) = (1+\alpha)B + \frac{1}{2}$. The indifference curve representing a particular value of $E(T|B)$ is a straight vertical line and the school would admit any students with a public endowment above the population median. This gives us the blue shaded area in Panel (b).

Putting test-required and test-blind in one graph, Panel (d) shows that test-blind benefits the students that have high public endowment but low private endowment (in the yellow shaded area), since they are able to hide their less-favorable standardized test scores behind their above-average public endowment. It hurts applicants with low public endowment but high private endowment (in the green shaded area) because it shuts down the channel (standardized test) through which they can stand out and showcase their competence.

Compared to test-required, test-blind reduces the academic preparedness and diversity of admitted students, if we define academic preparedness as the average A and diversity as the standard deviation of B of the admitted class. To see this, imagine we replace a random student in the green area with a random student in the yellow area. It is easy to show that the replaced student always have a higher $A = \alpha B + e$ and a lower B than the replacing student, which will pull down the average academic preparedness and diversity of the admitted class. This implies that test-required would Pareto dominate test-blind if the school appreciates both academic preparedness and diversity.

Test-optional lies somewhere in between test-required and test-blind. As shown in Panel (c), when the school interprets non-reported A as $\bar{A} = \alpha B + c$, any student with $e > c$ would report A and any student with $e < c$ would not report. This implies that the school's expected total endowment for a student depends on her reporting decision R, namely $E(T|B,R) = E((1+\alpha)B + e|B,R) = (1+\alpha)B + max(e,c)$. This means the T-inference curve is kinked at $e = c$, with a downward slope of $-(1+\alpha)$ when $e > c$ and a slope of infinity when $e < c$. Again, the school looks for a T-indifference curve that corresponds to the population median and admits all students with the expected total endowment above it. This corresponds to the blue shaded area in Panel (c).

Similar to the case of test-required, the school rejects every student with $B < \frac{1}{2} - \frac{1-c^2}{2(1+\alpha)}$ no matter whether the student reports or does not report A (straight reject), admits every student with $B > \frac{1}{2} + \frac{(1-c)^2}{2(1+\alpha)}$ (straight accept), and trade offs between B and the reported A for any student in between ("tradeoff group"). Any student in this tradeoff group but does not report A would be rejected. Note that these cutoffs depend on c: when the school adopts a more lenient interpretation of non-reported score (higher c), it expands the straight reject and straight accept groups. As a result, the tradeoff group shrinks and fewer students choose to report, both of which diminish the information value of test score. When c is extremely lenient ($c = 1$), it eliminates the tradeoff group and the regime is equivalent to test-blind. When c is extremely harsh ($c = 0$), it motivates every student to report and the complete unraveling makes the regime equivalent to test-required.

Panel (e) of Figure 1 further compares the discrepancy of admission outcomes between test-required and test-optional. Similar to test-blind, test-optional rejects the low-SES-high-achieving students in the green area and admits the high-SES-low-achieving students in the yellow area. It is easy to show that replacing a random student in the green area with a random student in the yellow area would lead to a strict decline of A and a strict increase of B, pulling down the average academic preparedness and diversity of the admitted class. In short, test-required Pareto dominates any test-optional policy if the school appreciates both academic preparedness and diversity.

Panel (f) compares two test-optional policies with different leniency. As discussed before, a harsher interpretation of non-reporting would motivate more students to report, which pushes the downward sloping boundary of the admitted group to the left among reported students and the vertical boundary of the admitted group to the right among non-reported students. Consequently, a harsher test-optional policy rejects the high-SES-low-achieving students in the yellow area in exchange for the low-SES-high-achieving students in the green area, which increases the average academic preparedness and diversity of the admitted class.

Overall, the illustrative model concludes that the perceived tradeoff between academic preparedness and diversity is non-existent in the simplified admission problem: when private and public endowments are positively correlated, the school's interpretation of non-reporting is known to students, and all students are rational in their strategic reporting behavior, the Pareto dominance follows the order of:

$$\text{Test-required} \gg \text{Harsh test-optional} \gg \text{Lenient test-optional} \gg \text{Test-blind}.$$

Is there any scenario where this order of Pareto dominance may break down if we change some assumptions in the illustrative model? The lab experiment presented below would relax the assumption of student rationality and perfect information. Here we briefly discuss how the model would change if the school has different preferences on the students' private and public endowments.

It is not difficult to show that, as long as the school has a positive marginal utility on A and B and treat the two as perfect substitutes, we can redefine one of the two endowments, re-scale the total endowment function and the positive correlation between the two endowments, and make it equivalent to the illustrative model. The model will change if the school has a negative marginal utility on B but a positive marginal utility on A.

This would introduce a tradeoff between (1) a positive preference on B because higher B implies higher A, and (2) a fundamental distaste on B. If we assume (1) dominates (2) so that the school still prefers to admit students with higher B if A is not observable, then test-optional (or test-blind) could increase the average academic preparedness of admitted students above that of test-required. This is because it replaces some low-SES-high-achieving students with some high-SES-low-achieving students but the definition of high-achieving is compromised due to the school's fundamental distaste of B. As before, test-optional (or test-blind) still reduces the diversity of admitted students as compared to test-required, so we may have a tradeoff between lower diversity and better academic preparedness by different admission policies. The rest of the paper ignores this theoretical possibility because it is unrealistic to assume a typical college in the US would have a fundamental distaste on other application materials such as GPA and extracurricular activities.

2.2 A College Application Problem Without Perfect Information

One key assumption that we make in the simplified problem is that students have perfect information on the school's interpretation of non-reported private endowment under a test-optional policy. This, however, is rarely the case in reality as colleges do not fully disclose that information. Thus, college applicants' actual reporting behavior and admission outcome may deviate from what one would predict under perfect information. To better reflect the reality, this subsection extends the simplified problem by allowing students to play the admission game in multiple rounds, to obtain their own admission outcome in each round, and to observe some statistics of the college's admission outcomes in the past. The extended model is outlined below.

School. In period t, each student has a public endowment, B_t, and a private endowment, A_t. The school admits students based on $T_t = \bar{A}_t + B_t$, where $\bar{A}_t$ is the school's belief of a student's private endowment. Let the school's belief of a student's non-reported private endowment be given by the function $g(B_t, x_{t-1}, y_{t-1})$, where x_{t-1} and y_{t-1} are the 25th and 75th percentile of A_{t-1} of those who reported it and got admitted in the previous period (mimicking the 25th and 75th score percentiles reported in the U.S. News). Specifically, define $g_t = g(B_t, x_{t-1}, y_{t-1})$ as:

$$g(B_t, x_{t-1}, y_{t-1}) = \begin{cases} B_t & \text{with prob. } \alpha \\ \gamma_0 + \gamma_1 x_{t-1} + \gamma_2 y_{t-1} & \text{with prob. } 1 - \alpha. \end{cases} \tag{5}$$

With probability α, the school interprets non-reported private endowment as perfectly identified by the student's public endowment; with probability $(1 - \alpha)$, the school interprets non-reported SAT/ACT scores as a linear function of admission statistics of the previous entering cohort. Here we use three parameters ($\gamma_0, \gamma_1, \gamma_2$) to describe the function g_t because the classical unraveling theory implies that the school should interpret all non-reported A as the worst possible outcome, but the school may deviate from the classical unraveling theory for ideology reasons. In particular, a generous interpretation of non-reported A may imply a a high γ_2 and a low γ_1, but a cynical interpretation of non-reporting may imply a high γ_1 and a low γ_2. As described later, our experiment introduces some variations in γs to represent different school interpretation of non-reporting.

Moreover, students do not have perfect information on the school-belief function g_t.

Assume a student's belief of g_t takes the form

$$\hat{g}(B_t, x_{t-1}, y_{t-1}) = \begin{cases} B_t & \text{with prob. } \hat{\alpha} \\ \hat{\gamma}_0 + \hat{\gamma}_1 x_{t-1} + \hat{\gamma}_2 y_{t-1} & \text{with prob. } 1 - \hat{\alpha} \end{cases} \tag{6}$$

so that a student's subjective expectation of the school's guess, $\bar{A}_t$, is:

$$\mathrm{E}_s[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] = \hat{\alpha} B_t + (1 - \hat{\alpha})[\hat{\gamma}_0 + \hat{\gamma}_1 x_{t-1} + \hat{\gamma}_2 y_{t-1}] \tag{7}$$

Student. Let p denote the probability of admission, and let U^a, U^r ($U^a > U^r$) denote the utilities from admission and rejection, respectively. A student's utility maximization problem in period t is: given B_t, $\max_{R_t \in \{0,1\}} \text{EU} = p_t U^a + (1 - p_t) U^r$. We can write p_t as a weakly increasing function f of total endowment, T_t. It follows that maximizing EU is equivalent to maximizing $p_t = f(T_t) = f(B_t + \bar{A}_t)$. Then, a student reports in period t if

$$\begin{aligned} \underbrace{f(B_t + A_t)}_{\text{Admission prob. from reporting}} &> \underbrace{\hat{\alpha} f(B_t + B_t) + (1 - \hat{\alpha}) f(B_t + \hat{\gamma}_0 + \hat{\gamma}_1 x_{t-1} + \hat{\gamma}_2 y_{t-1})}_{\text{Expected admission prob. from non-reporting}} \\ &= f(B_t + \mathrm{E}_s[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] - \pi(\rho)) \end{aligned} \tag{8}$$

where $\pi(\rho)$ is an increasing function of some measure of risk aversion. The second equality gives the value of total endowment such that the student would be indifferent between having that value and a lottery indicated by the RHS of the first inequality. Since $f(T_t)$ is increasing in T_t, the decision rule is simplified to:

$$R_t = \mathbb{1}\Big\{ A_t > \mathrm{E}_s[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] - \pi(\rho) \Big\} \tag{9}$$

Note that $\pi(\rho)$ does not represent aversion to the admission probability. It can be interpreted as the aversion to the uncertainty in student's mind about how the school would interpret non-reporting (i.e. the uncertainty about the true α and γs). Note that the uncertainty may still exist even if the students chooses to report, because different school interpretation of non-reporting may affect her relative ranking among all applicants. Holding everything else constant, more risk averse players may be more or less likely to report their private endowment, A_t, depending on how she perceives the uncertainty would affect her differently in the case of reporting versus non-reporting.

3 Experiment Design

In our experiments, subjects completed 50 rounds of game and additional tasks depending on the treatment. Instructions of the experiment were presented and read to the subjects at the beginning of the session. The Appendix contains the full instructions. At the end of each session, subjects were paid, privately and in cash, their show-up fee plus any additional earnings from the experiment.

Our main sessions were conducted at the Experimental Economics Lab at the University of Maryland (EEL-UMD). [5] In this laboratory, subjects were separated with dividers, and each subject was provided with a personal computer terminal.

Each Round. In our experiments, the subject was the sender (or the "applicant") and the computer was the receiver (or the "program"). In each round and for each player, the computer randomly drew a whole number from the set $\{1, 3, 5\}$, called the "public endowment". Each number in the set was equally likely to be drawn. Then, for each subject, the computer drew a second whole number, called the "private endowment", with the following rule: with 50% chance the private endowment is equal to the public endowment, with 50% chance the private endowment is chosen from the set $\{1, 2, 3, 4, 5\}$ with equal probability on each number in the set. For example, if the randomly chosen public endowment was 3, then with 60% chance the private endowment would be 3, and each number in the set $\{1, 2, 4, 5\}$ has 10% chance of being chosen. The rule was designed such that the correlation between the public and private endowment was around 0.5, a number close to the reported correlation between high school GPA and SAT scores (Westrick et al., 2020).

Each subject was shown her public endowment and private endowment. Subjects were made aware of the state spaces of both endowments and the positive correlation between the two, but were not told the probability distribution of either endowment. Each subject was also shown the admission statistics from the previous round: the 25th and 75th percentiles of the private endowment for those who reported it and were subsequently admitted, and the mean of the public endowment for those who were admitted. Then, subjects were given the option to either "report" or "not report" their private endowments, with the understanding that the computer knew their public endowments. There was no time limit for the subject decision.

After all subjects made their reporting decisions, the computer calculated a total endowment for each subject. If a subject's private endowment was reported, the total endowment would be the sum of the actual public and private endowments. If a subject's

[5] Our experiment was programmed and run using oTree (Chen, Schonger and Wickens, 2016).

private endowment was not reported, the computer took a guess on it, and assigned the sum of public endowment and the guess of private endowment as the total endowment. Then, the computer ranked all subjects in the session by their total endowments, and admitted the top half of subjects. Subjects were shown their own admission results and the admission statistics of the current round, which would be reminded in the next round.

At the end of the experiment, two random rounds were selected for payment. Each subject were paid $6 for each admission in those rounds. The maximum amount of payment through this channel is $12, when the subject was admitted in both randomly selected rounds. Therefore, it is in each subject's best interest to be admitted by the program in every round.

Treatment Variation. Our primary treatment variations occurred on the program's guess of a subject's private endowment if it was not reported. In other words, the treatment variations came from the selection of parameters in the expression of school-belief function $g(B_t, x_{t-1}, y_{t-1})$ in Equation 5. [6]

T1: $g_t = B_t$ with prob. 0.5, $g_t = 0.5x_{t-1}$ with prob. 0.5

T2: $g_t = B_t$ with prob. 0.5, $g_t = 0.5x_{t-1} + 0.5y_{t-1}$ with prob. 0.5

Our main sessions only varied γ_2, while setting $\gamma_0 = 0$, $\gamma_1 = 0.5$, and $\alpha = 0.5$ in both the real correlation of public and private endowments and the weight that the school puts on public endowment when it takes a guess about non-reported private endowment. In the first treatment (T1), we set $\gamma_2 = 0$; in the second treatment (T2), we set $\gamma_2 = 0.5$. Therefore, the program places harsher punishment on non-reporting in T1 than in T2. All subjects in a session were randomly assigned to a treatment for the entire session at the beginning of the experiment. Subjects in different treatments were given the same instructions, but might receive differential feedback through their own admission outcomes and the admission statistics in each round.

After all subjects completed 50 rounds, we used the multiple price list method (Holt and Laury, 2002) to elicit risk attitudes. This allows us to control for subjects' relative risk preferences when modeling their reporting choices. We randomly selected one lottery choice in each session and paid the subjects accordingly. The complete list of of lottery choices are included in the Appendix.

[6] In our "independent treatment", which is not discussed in this draft, we have additional treatment variations. Specifically, we set $\alpha = 0$, $\gamma_1 = 0$, and vary γ_2.

4 Results

In our main sessions, we observed 285 subjects making a total of 14,250 reporting decisions. Over 18 sessions, the mean session size was approximately 16. We used a show-up fee of $10, and on average subjects earned $18.30. The minimum payment was $11, and the maximum was $27.

We assigned 144 subjects to the first treatment (T1), and 141 subjects to the second treatment (T2). Table 1 shows the summary statistics of subjects in the main sessions. We had a slightly higher number of women than men in both treatments. All subjects were undergraduate students, with roughly half of them being white, and 40 percent of them being freshmen. Subjects' experience from college applications and standardized tests were balanced across treatments. They had similar number of test attempts, SAT and ACT test scores, number of schools applied. They also made similar SAT and ACT submission choices when they applied to college.

To complement our lab results, we also run simulations under the same setting (16 players per session, 9 sessions each for T1 and T2) but where players have perfect information on the school-belief function g_t and report if their true private endowment is higher than the expected school-belief of their non-reported private endowment, i.e.

$$R_t = \mathbb{1}\Big\{A_t > \mathrm{E}[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0]\Big\} \tag{10}$$

Note that the expectation here does not have a subscript s due to its objectivity. Real and simulated subject behavior may differ in that: (i) real subjects do not know g_t but the computer knows, (ii) real subjects may be risk averse or risk loving but the computer is not, (iii) real subjects may not be full rational but the computer is, and (iv) real subjects may learn between rounds but the computer does not. In our structural model, we address (i) by estimating subject belief of g_t or eliciting it from survey questions, (ii) by including a risk measure of each subject, (iii) by introducing a probability of irrationality, and (iv) by only focusing on later-round results.

4.1 Admission Statistics

Figure 2 shows the trends in admission statistics, i.e. the 25th and 75th percentiles of the private endowment of those who reported it and were offered admission in the previous period. Panels (a) and (c) come from lab data under T1 and T2, respectively, and (b) and (d) come from the corresponding simulation results.

In the first round, we set the initial "previous" 25th and 75th percentiles of the private endowment as 2 and 4. The 75th percentiles quickly converge to 5 in both treatments and in both simulations. For T1, the 25th percentile goes up immediately and starts fluctuating around 4. A similar trend is observed in our simulation. For T2, we see a small and gradual increase in the 25th percentile of private endowment after the initial round and till round 10. It then starts fluctuating around 4.5, being slightly lower than the simulated 25th percentile in most rounds. In general, we see that the admission statistics in the lab are similar to those from the simulation, indicating only small differences in the composition of the entering cohort between the field and the world where students are well-aware of the school's belief and behave rationally. Given a pre-determined school interpretation of non-reporting, the convergence of admission statistics is fast, even with a small sample per application cycle. The difference in admission statistics between T1 and T2 also signifies subjects' ability to learn from the feedback available to them. Since subjects in T1 and T2 were provided with identical instructions at the beginning of the experiments, the difference in learning experience came from the admission statistics and their own admission outcomes from previous rounds.

4.2 Reporting Decision

Panel A of Table 2 presents the average reporting rate in T1 and T2 by subjects' private endowment. As predicted by the "partial unraveling" incentive, subjects with higher private endowments are more likely to report: in T1, the reporting rate increases from 21.7% when $A = 1$, to 29.4% if $A = 2$, 49% if $A = 3$, 81.8% if $A = 4$, and 93.5% if $A = 5$. This monotonic relationship between reporting and private endowment continues to hold in T2, but the absolute magnitudes of reporting rate decline from T1 to T2 for every level of private endowment except for $A = 5$. Again, this is consistent with the partial unraveling incentive because the school in T2 is more lenient in its interpretation of non-reporting.

If we compare the starting rounds (1-20) and ending rounds (21-50) within each treatment, the reporting rate declines over time in T2 for all $A < 4$, but it only drops in T1 for $A = 1$ and $A = 2$. Since subjects started with exactly the same setting in T1 and T2, this suggests that subjects' initial belief of school interpretation of non-reporting may be closer to the true interpretation of T1 than to that of T2. Over time, subjects learn that the school in T2 is more lenient and therefore become more likely to withhold their low scores.

Panel B of Table 2 summarizes the average reporting rate in T1 and T2 by subjects' public endowment. In both T1 and T2, the reporting rate is slightly lower when $B = 3$

than when $B = 1$. This is consistent with the reverse unraveling incentive, because by definition subjects with $B = 3$ are more likely to receive higher A than subjects with $B = 1$ but they are no more likely to report. However, when the public endowment increases to $B = 5$, the reporting rate is 66.2% in T1 and 64.7% in T2, which seems much higher than the reporting rate when $B = 1$ or $B = 3$ (38-50%). At the first glance, this pattern goes against the reverse unraveling prediction. This is because subjects with a higher B are also more likely to receive a higher A, and the average reporting rate by B mixes the partial unraveling incentive with the reverse unraveling incentive.

More specifically, Figure 3 presents the observed reporting rates given the public endowment B and private endowment A. Here we compute the reporting rates for rounds 21 to 50 because over 80 percent of subjects report that they have formed a belief of the school's interpretation of non-reporting by round 20. [7]

If we focus on the same private endowment (say $A = 3$), most subjects with the lowest public endowment ($B = 1$) are eager to report (89.6% in T1 and 81.9% in T2), but most subjects with the highest public endowment ($B = 5$) are reluctant to report (only 21.4% report in T1 and 11.9% report in T2). The same pattern occurs for $A = 1$ and $A = 2$. Such a drastic decline of reporting rate by B, conditional on the same A, reflects the reverse unraveling incentive.

If reverse unraveling holds, how can we explain the relatively high average reporting rate in Panel B of Table 2 for all subjects with $B = 5$? The main reason is that when public endowment was the highest ($B = 5$), by construction the private endowment would also be the highest ($A = 5$) with 60% chance, and over 90% of the highest private endowments were reported during the experiment because of the partial unraveling incentive. Thus, the high average reporting rate by B, which is unconditional on A, reflects a mixture of the reverse unraveling incentive (conditional on the same A) and the partial unraveling incentive (across different A).

Another way to read Figure 3 is to compare subjects' reporting behavior with what we would expect in theory if subjects are fully rational and know the school's interpretation function beforehand. Table 3 shows the hypothetical reporting decision indicated by Equation 10. If subjects were successful in learning from their admission results in the first 20 rounds and forming an accurate belief of the school-belief function g_t, we should only observe significant gaps in reporting rates between T1 and T2 under the following three cases: ($A = 2$, $B = 1$), ($A = 3$, $B = 3$), ($A = 4$, $B = 5$). To be specific, in these cases a perfect-information fully-rational subject would always report under T1 and always not report under T2 because the school is more lenient for non-reporting in T2.

[7] Figure A.3 shows the rounds at which subjects formed the final belief of the school's interpretation.

Figure 3 shows a few deviations from these hypotheses. When $B = 1$, we observe no significant difference in reporting rates between T1 and T2 at $A = 2$ (mainly due to over-reporting under T2), but significantly higher reporting rate under T1 at $A = 1$. When $B = 3$, we do observe significantly higher reporting rate under T1 at $A = 3$, but it is still far from full reporting (at 49%). When $B = 5$, we observe higher reporting rates under T1 at $A = 4$, but the difference is not statistically significant (mainly due to over-reporting under T2). There is also significantly higher reporting rate under T1 at $A = 3$. These deviations signal differences between the hypothetical decision rule given by Equation 10 and the actual decision rule given by Equation 9, suggesting that lab subjects may have some departure from full rationality or perfect information.

To better understand subjects' round-by-round learning, Figure 4 reports the trends in subject reporting rates from round 1 to round 50 in T1 and T2 separately. We do not see a clear trend or convergence in reporting rates under T1: the lines are fairly flat over time regardless of the public endowment. Similar (absence of) trends are observed under T2 when public endowment was either low ($B = 1$) or high ($B = 5$). When the public endowment was of medium value ($B = 3$), the reporting rate dropped from roughly 55% to 40% and then remained flat for the rest of the session. Panel (c) of Figure 4 highlights the difference in reporting rates between T1 and T2 over time. For subjects with medium public endowment, the reporting trend indicates that those in T2 may have learned from their own experience from previous rounds and realized a relatively lenient school policy. Although both treatment groups had similar reporting rates at the beginning of the experiment, the gap in reporting rates widened as they played more rounds. However, as we have seen in Figure 3, the reporting rates in both treatments are far from the simulated rates with perfect-information fully-rational subjects. In general, these evidence suggests limited subject learning but it has not yet associated a subject' own experience from previous rounds with the same subject's reporting decision in later rounds, which we will further explore in the next section.

Finally, Figure 5 shows the reporting rates by gender and race in rounds 21 through 50. The reporting rates are significantly higher for male in T1, but not in T2. When we look at reporting rates by race, Asian subjects have roughly 5 percent higher reporting rates than Black and White subjects in T1. In T2, all three races have similar reporting rates. In both illustrations, the reporting rates are higher in T1 than in T2, suggesting that all subgroups realized that the school has a harsher interpretation of non-reporting in T1.

5 Structural Estimation of Reporting Decision

To study the relationship between the real school belief and subject perception of it, we estimate a structural model of subject reporting decision. Recall that we assume a subject's belief of the school belief of non-reported private endowment, $\hat{g}(B_t, x_{t-1}, y_{t-1})$, is B_t with probability $\hat{\alpha}$ and $\hat{\gamma}_0 + \hat{\gamma}_1 x_{t-1} + \hat{\gamma}_2 y_{t-1}$ with probability $1 - \hat{\alpha}$. Consider a discrete choice model. Let the utilities of report and not-report in period t be given by

$$\begin{aligned} V_{R_t=1} &= A_t \\ V_{R_t=0} &= E_s(\bar{A}_t) - \pi(\rho) + \epsilon \\ &= \hat{\alpha} \cdot B_t + (1-\hat{\alpha})[\hat{\gamma}_0 + \hat{\gamma}_1 \cdot x_{t-1} + \hat{\gamma}_2 \cdot y_{t-1}] + \Gamma \cdot X + \epsilon \end{aligned} \tag{11}$$

where X is a vector of subject covariates and ϵ is the logit error. Given the utilities, a rational subject would report in period t if $V_{R_t=1} > V_{R_t=0}$.

Since subjects in different treatments were given the exact same instructions, subject learning was the major distinction of each treatment specification. We use two distinct models to capture subject learning at any point in the experiment: a Heuristic model and a Bayesian updating model.

Heuristic model. We define a total of eight learning variables for each subject at each round: subject reporting rate under the same private endowment in all previous rounds, subject reporting rate under the same public endowment in all previous rounds, subject acceptance rate pooled and by reporting decision under the same private endowment in all previous rounds, subject acceptance rate pooled and by reporting decision under the same public endowment in all previous rounds. These variables provide a simple but comprehensive summary of a subject's experience during the experiment up to a given round.

Bayesian updating model. Suppose at the beginning of round 1, each subject starts with a prior that α can take k_α values between 0 and 1 with a density distribution $\phi_\alpha^{(1)}$, and similarly for γ_0 and (γ_1, γ_2). For simplicity, let $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$, $\gamma_0 \in \{0, 1, 2, 3, 4\}$, and $(\gamma_1, \gamma_2) \in \{(0,0), (0,0.5), (0,1), (0.5,0), (1,0), (0.5,0.5)\}$. We refer to these priors as $\phi^{(1)} = \{\phi_\alpha^{(1)}, \phi_{\gamma_0}^{(1)}, \phi_{\gamma_1,\gamma_2}^{(1)}\}$. This gives us 150 possible combinations of α, γ_0, γ_1, γ_2, which represent 150 unique school admission policies.

Assuming the prior densities of $\{\alpha, \gamma_0, \gamma_1, \gamma_2\}$ are independent of each other, we can also express the prior as a vector of 150 probabilities corresponding to each of the 150 school policies, namely $\phi^{(1)} = \{\phi_k^{(1)}\}$ where $k = 1, .., 150$. In each round, the subject compares the utility of reporting vs. non-reporting in Equation 11 to make her optimal decision of reporting (note that in round 1, there is no previous round so we set x_0 and y_0 to be 2 and 4, respectively). For any rational subject in round t, given her endowments (A_t, B_t), reporting decision (R_t), admission outcome (O_t), and class statistics (x_{t-1} and y_{t-1}), she updates her belief from $\phi^{(t)} = \{\phi_k^{(t)}\}$ to $\phi^{(t+1)} = \{\phi_k^{(t+1)}\}$ by the following:

$$\begin{aligned}\phi_k^{(t+1)} &= \Pr(k|A_t, B_t, R_t, x_{t-1}, y_{t-1}, O_t, \phi^{(t)}) \\ &= \frac{\Pr(O_t|A_t, B_t, R_t, x_{t-1}, y_{t-1}, k) \cdot \phi_k^{(t)}}{\sum_{k'=1}^{k'=150} \Pr(O_t|A_t, B_t, R_t, x_{t-1}, y_{t-1}, k') \cdot \phi_{k'}^{(t)}}\end{aligned} \tag{12}$$

To compute the probability of receiving admission outcome O_t, we assume a focal subject believes every other subject follows exactly the same belief as the focal subject. Then, we simulate each conditional probability and derive the posterior belief $\phi^{(t+1)}$, which will be a function of the prior $\phi^{(1)}$, and the focal subject's reporting history $\{R_1, R_2, ...R_t\}$, admission history $\{O_1, O_2, ...O_t\}$, and class statistics $\{x_1, x_2, ..., x_t, y_1, y_2, ..., y_t\}$. This gives us the likelihood of reporting for each subject at each round. Finally, we estimate the parameters $\{\phi^{(1)}, \Gamma\}$ by maximum likelihood. The detailed simulation and estimation procedures are provided in the appendix.

Estimation results. The results from the heuristic model are provided in Table 4. We set the coefficient of A_t in the logit model to always be 1, so that the coefficients in the table can be translated to the model parameters with a minus sign. Columns 1 and 2 show that limiting the sample to the last 30 rounds do not change the estimates of the logit model. In other words, there is no direct evidence for subjects updating their reporting decision rule overtime, given their private endowment, public endowment, previous admission statistics, and risk attitudes. However, when we include the learning variables and a full set of subject covariates in the model, the coefficient for public endowment more than doubled, while the coefficient for admission statistics from the previous round decreased in magnitudes. [8] The estimates in the first two columns are likely biased as subject demographics and their previous experience both in college application and in the lab influenced their reporting decision rule. In the last column we add subject fixed effects to the model. The coefficient for the 25th percentile of private endowment decreased by half and become insignificant, while the others are similar to those in the third column (hereinafter "Model 3").

In both treatments in the lab, we assign 50% weight to the public endowment in the school's interpretation of non-reported private endowment (i.e., $\alpha = 0.5$). The coefficients in the last two columns suggest that the elicited weight on public endowment are 0.72-

[8] The set of subject covariates include subject demographics and their own experiences in college application, e.g., standardized test scores, reporting decisions. They are summarized in Table 1.

0.75. Thus, subjects tend to overweight their public endowment and underweight school attributes when they make reporting decisions.

In Table A1 we report the estimated $\{\phi^{(1)}, \Gamma\}$ from the Bayesian updating model. We define the average prior and average posterior as the "average school policy" that subjects had in mind in round 1 and round 50, respectively. In other words, we compute the average α, γ_0, γ_1, γ_2 from the estimated parameters. There is one average prior for all subjects and one average posterior for each treatment group. Given the estimates, the average prior refers to school policy where $\mathrm{E}[\bar{A}_t|B_t, x_{t-1}, y_{t-1}, R_t = 0] = 0.428B_t + 0.575 + 0.173x_{t-1} + 0.076y_{t-1}$. The average posterior is $0.478B_t + 0.685 + 0.122x_{t-1} + 0.076y_{t-1}$ for T1 and $0.514B_t + 0.703 + 0.100x_{t-1} + 0.103y_{t-1}$ for T2. These suggest that: first, previous admission statistics (x_{t-1} and y_{t-1}) play a smaller role in subjects' belief of the school policy, compared to the true school policy. Second, consistent with what we have observed earlier, subjects in T1 recognized that the computer had a harsher interpretation on the non-reported private endowment compared to subjects in T2. If we plug in the actual average admission statistics (x_{50}, y_{50}) in our T1 and T2 sessions, the average posterior is $0.478B_t + 1.591$ for T1 and $0.514B_t + 1.692$ for T2. [9] Note that plugging in the same admission statistics to the true school interpretation in Table 3, we get $0.5B_t + 1.078$ for T1 and $0.5B_t + 2.435$ for T2. Thus, while the T1 posterior is smaller than the T2 posterior, indicating that subjects learn the crucial difference between T1 and T2 and choose their strategic reporting decision accordingly, neither of the average posteriors is exactly equal to the accurate school belief. Consistent with the observed reporting rates (for example, shown in Figure 4 Panel c), T1 posterior is higher than the true T1 school belief and T2 posterior is lower than the true T2 school belief.

The predicted subject reporting rates from the Heuristic model are presented in Table 5, Panel A. Models 3 and 4 predict reporting rates much closer to the actual rates than Model 2, suggesting that an model of subject reporting decision should include subject attributes and subject learning. Adding subject attributes or subject fixed effects also substantially increases the model's likelihood: from -4,018 to -3,029 or -2,835. While the average distance between pooled actual and predicted reporting rates are low (1.5 for Model 3), the difference could be quite large when we focus on a particular treatment. In particular, the actual reporting rate in T2 when the private endowment was 3 was 30.4%, while the model prediction is 35.1%; the actual reporting rate in T2 when the private endowment was 4 was 80.5%, while the model predictions are 76.6%. In general, the model fit is better in T1 than in T2.

Figure 6 presents the actual and model predicted reporting rates for both the Heuristic

[9] The average x_{50} is 4.31 for T1 and 4.74 for T2. The average y_{50} is 5 for both T1 and T2.

model (Model 3) and the Bayesian updating model. The Heuristic model has a much better model fit than the Bayesian updating model. While the average distance between the actual and predicted reporting rates is 1.5% for the Heuristic model, the average distance is 4.7% for the Bayesian updating model. Therefore, for the remainder of the paper, we focus on the Heuristic model due to its accuracy and simplicity.

Subject Naivety. To incorporate the possibility that a subject may naively follow some rule of thumb and do not engage in any learning from round to round, we introduce two naivety parameters that capture the likelihood with which the subject chooses report or not report naively in each round. We assume the naive rule of thumbs are:

$$\text{Rule of thumb 1:} \quad R_t|\text{naive} = 1 \text{ if } A_t \in \{4, 5\} \tag{13}$$

$$\text{Rule of thumb 2:} \quad R_t|\text{naive} = 0 \text{ if } A_t \in \{1, 2, 3\} \tag{14}$$

Let the probability of being naive be θ_1 when the private endowment was less than or equal to 3, and θ_2 when the private endowment was larger than 3, then the likelihood of reporting is

$$\Pr(R_t = 1) = \begin{cases} (1-\theta_1) \cdot \Pr(R_t = 1|\text{rational}) & \text{if } A_t \in \{1, 2, 3\} \\ \theta_2 + (1-\theta_2) \cdot \Pr(R_t = 1|\text{rational}) & \text{if } A_t \in \{4, 5\} \end{cases} \tag{15}$$

where $\Pr(R_t = 1|\text{rational}) = \frac{\exp(V_{R_t=1})}{\exp(V_{R_t=1})+\exp(V_{R_t=0})}$. The predictions of our model with naivety are presented in Table 5, Panel B. The addition of naivety parameters increases total log likelihoods, and reduces the average distance between pooled actual reporting rates and predicted reporting rates from a range of 1.4-6.8 to a range of 1.0-3.1. In Model 3, we estimate θ_1 to be 3% and θ_2 to be 7%, which are small but non-trivial. The average distances are reduced by roughly 20% from those in Panel A. When we look at predictions in a particular treatment, the average distance between the actual and predicted rates becomes slightly larger in T1 and smaller in T2. In general, the model's ability to explain subject reporting decision is better when we include naivety parameters.

6 Trading Off Academic Preparedness with Diversity

In this section, we first discuss the tradeoff between academic preparedness and diversity of a school's entering cohort in our stylized model. Then, we extend our stylized college application problem and discuss how the preparedness-versus-diversity tradeoff may change if we allow for applicant-group admission quota, resource constraints in test-taking, or an additional dimension of student application profile, respectively.

6.1 Tradeoff in the Stylized Model

Assuming that schools always admit the top candidates (i.e. candidates with the highest overall endowment) until they reach full capacity, the composition of the admission cohort will be different when schools commit to different interpretations of non-reporting. At one extreme, a school that mandates standardized test reporting aims at admitting students with the highest academic preparedness, but some may worry that this is at the expense of a less diverse entering cohort, offering few opportunities for disadvantaged students with little test-preparation resource. At the other extreme, a school that does not consider standardized test as part of the college application intends to attract a more diverse application pool, but may have limited ability to identify students with the best school readiness. Most schools that adopt a test-optional policy fall between these two extremes. While these schools have their own objective on the composition of the entering cohort, it is still unclear how the tradeoff between academic preparedness and diversity looks like given a school's interpretation of non-reported standardized test scores.

To demonstrate this tradeoff for school policies other than those appeared in our experiments, we use the structural model in Section 5 to simulate student reporting decisions and admission outcomes for counterfactual school interpretations.

Given any set of $\{\alpha, \gamma_0, \gamma_1, \gamma_2\}$, we run the simulation as follows: first, we simulate a pool of 16 subjects with each subject being a "representative" subject in our experiments. In other words, for categorical subject attributes (e.g., gender, race, school year), the simulated subject pool will have the same composition as those in our lab sessions; for non-categorical subject attributes (e.g, SAT/ACT scores, number of SAT/ACT attempts, number of schools applied), we assign each simulated subject the average value of those in our lab sessions. Second, we assign private and public endowments under the same procedure as in the lab. We construct the learning variables defined in Section 5, and update them for every round of simulation. We simulate 50 rounds of reporting decisions and admission outcomes given the school interpretation, the estimated model parameters from Table 4, Model 3, and the naivety probabilities reported in Table 5. Finally, we repeat the first two steps for each counterfactual setting for 100 times. Keeping the subject size and the number of rounds consistent to those in the lab will allow us to compare the simulated results to our experiment results, and running a large number of simulated sessions will absorb the variation in the simulated results caused by the small sample size in each session. We use the average outcomes in rounds 35-50 as bases for computation in this section.

Table 6 lists the two school interpretations we have used in the experiments (T1 and T2), along with 16 alternative school interpretations (C1-C16) we will try in counterfactual simulations. For example, C1 assumes that the school would interpret a non-reported private endowment as a simple average of the subject's public endowment and the 25th percentile of the reported private endowment in the last admitted cohort. This is more pessimistic than C2, where the school uses the 75th percentile instead of the 25th percentile from the last admitted cohort, but more optimistic than C3, where the school uses the lowest possible private endowment (1) instead of 25th or 75th percentile when computing the simple average. C8 is even more pessimistic than C3, by putting more weight (75% instead of 50%) on the lowest possible private endowment and less weight (25%) on the subject's public endowment. In contrast, C10 is even more optimistic than C2, as the school's non-report interpretation put 75% weight on the highest possible private endowment. This is equivalent to assuming a non-report student will receive the highest test score with 75% of probability and receive the same score as her public endowment in the remaining 25% probability. Another notable counterfactual interpretation is C16, where the school simply assumes a subject's non-reported private endowment equal to her public endowment. This is different from test blind though, because the counterfactual belief only applies when the subject chooses not to report her private endowment and if the student is rational, the strategic choice of non-reporting would only occur if her private endowment is no better than her public endowment. Overall, the 16 counterfactual situations are designed to capture various school interpretations of non-report, which allow us to simulate how students may react strategically to these school interpretations. Departing from typical simulations of market equilibrium, we do *not* require the school's interpretation to align with the realized score distribution of non-reporting candidates, as schools may choose certain interpretation for ideological or other reasons.

To characterize the tradeoff between academic preparedness and diversity, we define "Academic Preparedness" as the average private endowment of the admitted cohort, and "Diversity" as the standard deviation of public endowment of the admitted cohort. We assume that, conditional on the quality of students that they admit, colleges prefer a more diverse cohort; and conditional on the diversity of the admitted students, collages prefer students with better academic preparedness. [10] In other words, assuming private endowment captures a student's standardized SAT/ACT test scores and public endowment

[10] In a robustness check, we use the proportion of lowest public endowment ($B = 1$) subjects being admitted as an alternative measure of diversity. The tradeoff using this measure is shown in Figure A.4. The two measures of diversity give nearly identical illustrations.

captures the student's non-test attribute, we consider college objective functions that are increasing in admitted student's SAT/ACT scores and in the dispersion of the non-test attribute, which presumably has a higher correlation with the applicant's socioeconomic status.

Figure 7 illustrates the tradeoff for the full set of school policies, including the two treatments in our experiment (T1, T2), test-required (TR), test-blind (TB), and the counterfactuals (C1-C16). There are a few main takeaways from this figure.

Most strikingly, there does not appear to be much of a tradeoff: school policies that admit students with higher academic preparedness also admit students from a more diverse background. In particular, test-required leads to an admission cohort with the highest average private endowment. It is clearly desirable if a school wants to prioritize the admission of students with better academic background. It also admits one of the most diverse cohorts in non-test attributes, because it provides students with a weak non-test attributes an opportunity to stand out through standardized tests. By contrast, with a test-blind policy, the school cannot distinguish applicants with high standardized test scores given the non-test attribute. Assuming a positive correlation between standardized test performance and the strength of non-test attribute, the school will admit those that have the highest public endowments. This reduces the diversity of the public endowment. Also, these students may or may not have high private endowments, thus the average private endowment is brought down. This explains why test-blind is dominated by all other policies plotted in Figure 7 in both academic preparedness and diversity. By ignoring an important information (standardized test), it ties the hands of a school, leaving less room for the school to find the best strategy to optimize its objective function.

Voluntary reporting raises a lesser degree of the same concern. Compared to a test-required policy, test-optional makes standardized test performance less visible, and forces the school to rely more on the public endowment for admission. This partially reduces the diversity of public endowment and the average of the private endowment. Empirically, 11 of the 16 counterfactual test-optional policies fell somewhere in between test-required and test-blind. They dominated test-blind but are dominated by test-required. Algorithms that give the most generous interpretation to non-reporting (e.g., C7, C10), and thus are closest to test-blind, are the ones that achieve the lowest academic preparedness and lowest diversity. For those that penalize non-reporting the most (e.g., C3, C8, C13), and thus are closest to test-required, opposite results are found as expected.

Interestingly, our simulations suggest that some test-optional policies that enforce severe punishment on non-reporting (e.g. C3, C8) may admit students from a more diverse non-test attributes than test-required. This would not be the case when applicants have perfect information on the school's interpretation of non-reported test scores.

Figure 8 presents the simulated tradeoff when students have perfect information on the school's back-end algorithm. This could happen when school publicly announces and credibly commits to its interpretation of non-reported standardized tests, and students fully understand the announced policy. In contrast, students in the real world may exhibit some naivety due to the lack of application information or the inability to fully comprehend a school's policy. In the imperfect-information simulations (Figure 7), we assume the same naivety probabilities as those reported in Table 5, Column 3. Comparing Figures 7 and 8, we observe that the simulated outcomes with perfect information are more aggregated and positioned very similarly as those with imperfect information, except that test-required would lead to a more diverse cohort under perfect information than all test-optional policies.

What drives the difference we observe (on C3 and C8 for example) between Figure 7 and 8? Using our lab setting as background, when the school punishes non-reporting really hard, in the perfect information case almost everybody would report. However, when subjects are not aware of the extremely harsh punishment, they don't report when they have relatively low or really low A compared to B because of the reverse unraveling incentive. For example, as shown in Figure 3, when $B = 3$, the reporting rates are low when $A <= 3$. This will provide some opportunity for low public endowment and high private endowment subjects (e.g., $B = 1, A = 5$) to get admitted because now they have a very good chance to win against some medium public endowment subjects (e.g., $B = 3, A = 3$) since the latter might not report. The admittance of low public endowment subjects then contributes to the increase in diversity in Figure 7. Similar reasoning can be applied to the real-world college application process, in which some low-SES-high-achieving students may benefit from (undisclosed) test-optional policies that severely punish non-reporting but some students with a relatively high income do not fully understand this and choose not to report due to reverse unraveling. Nevertheless, for most other test-optional policies, fully disclosing or hiding the actual school interpretation of SAT or ACT scores if they are not reported leads to similar admission portfolios.

To summarize, our simulations show that a test-blind policy is dominated by either test-optional or test-required policy in the academic preparedness and the diversity of the admission cohort. Test-required admits students with the highest academic preparedness and from a diverse background. While most test-optional policies are dominated by test-required, some may be desirable when the school prioritizes the diversity of its entering cohort and severely punishes SAT/ACT non-reporting. However, this gain of diversity (at some cost of academic preparedness) is likely transitory as we only find it present when students are not fully aware of the school's admission policy.

6.2 Extension #1: Imposing Admission Quota by Public Endowment

The results shown in Figure 7 are consistent with what we have illustrated in Section 2 for the simple college application problem (Figure 1). Test-required benefits applicants with low non-test attribute and high standardized test score, thus allowing a more diverse admission cohort. Test-optional admits more high non-test attribute applicants who hide their unfavorable standardized tests. The more generous the school is with regard to non-reporting, the more low non-test attribute, high standardized test score applicants would be hurt (by not getting admission). Test-blind allows the least diverse admission cohort because it only admits applicants with the highest non-test attributes.

In reality, schools may refrain from admitting students with the highest perceived total endowment as some argue they may not be the ones with the highest marginal returns from schooling (Dale and Krueger, 2002; Brand and Xie, 2010). Figure 9 shows otherwise identical compositions as in Figure 1 but with the school imposing an admission quota by public endowment. In particular, we assume the school categorizes applicants into three categories: applicants with high, medium, or low public endowments. The school admits one third of their students from each level. We see the exact same patterns as those in Figure 1. The previous conclusions hold within each group of applicants. Within each level of public endowment, test-required admits the most applicants with low public endowments, test-blind admits the least, and test-optional lies between the two.

6.3 Extension #2: Resource Constraints in Test-Taking

Students from sophisticated families with well-educated parents often are better-prepared for college applications. One major difference between these students and students from disadvantaged families is the probability of taking standardized tests prior to application. In 2023, the number of students from the top two quintiles of the family income distribution who take SAT tests is roughly twice that of students from the bottom two quintiles. [11] The lack of resource for test-preparation and test-taking puts students from low-income families at a further disadvantage under a test-required policy. Thus, a test-optional or test-blind admission policy may attract more low-income students to submit applications and potentially admit a higher proportion of low-income high-achieving students.

[11] Source: https://reports.collegeboard.org/media/pdf/2023-total-group-sat-suite-of-assessments-annual-report\%20ADA.pdf.

Under our setting, we simulate the reporting decision and admission outcome with heterogeneous test-taking. Let the probability of taking SAT/ACT tests be p_1, p_3, and p_5 for students with public endowments 1, 3, and 5, respectively. Suppose each applicant has a private endowment and the SAT/ACT scores perfectly identify the private endowment of an applicant if she takes the standardized test. We assume the probability of test taking is the same for applicants with the same public endowment, regardless of their private endowment. Under test-blind and test-optional, we assume all applicants apply for the school, and the school is not able to distinguish those who took the test and those who did not, conditional on not observing a test-score. Under test-required, only students who took the test would apply. We simulate a total of 9 sessions with a sample of 300 applicants (100 applicants for each public endowment) in each session.

Panel (a) of Figure 10 shows how the academic preparedness and diversity measures change as the probability of test-taking changes. The test-taking probabilities are: $p_1 = 0.5$, $p_3 = 0.75$, $p_5 = 1$ for TR_h1; $p_1 = 0.33$, $p_3 = 0.67$, $p_5 = 1$ for TR_h2; and $p_1 = 0$, $p_3 = 0.5$, $p_5 = 1$ for TR_h3. The figure suggests that, unless the probability of test-taking is extremely low for low public endowment applicants (e.g., none of the $B = 1$ applicants take tests), a test-required policy still dominates test-blind. In other words, the "benefit" to disadvantaged applicants from allowing them to apply a larger portfolio of schools, through not requiring standardized test scores, is hardly large enough to offset the "harm" from depriving their opportunity to stand out through standardized tests.

Panels (b)-(d) of Figure 10 present the academic preparedness and diversity tradeoff with different sets of p_1, p_3, and p_5. Test-blind is not affected because standardized tests are not required regardless of applicants' test-taking decisions. For counterfactual test-optional policies, each applicant's reporting decision and admission results are simulated based on the model derived in the previous section. In Panel (b) and Panel (c), the patterns are similar to those in Figure 7: test-required dominates test-blind, while some test-optional policies lead to a more diversified admission cohort than test-required. When none of the $B = 1$ applicants and only 50% of $B = 3$ applicants take tests, i.e. in Panel (d), test-blind and most test-optional policies result in more diversity than test-required. When applicants with low public endowments rarely take standardized tests, it mechanically limits test-required's ability to receive applications from diverse background and hence it is not surprising to observe some test-optional policies to achieve better diversity. In all cases, test-required admits the most academically prepared cohort and test-blind is dominated by either test-blind or test-optional policy. These figures send a clear message: test-required is the best policy if the school prioritizes academic merit, some test-optional policies may be favorable if the school puts great emphasis on a diverse entering cohort, and test-blind is never the optimal choice. As a robustness check, Figure 11 shows the tradeoff with heterogeneous test-taking behavior when subjects have perfect information on the school's interpretation and act rationally accordingly. The implications are very similar.

During college applications, minorities and students from low-income families are often disadvantaged in many ways other than the ability to take standardized tests: they have less test-preparation resources so the SAT/ACT scores may not reflect their true merit, they may have a higher cost of taking the tests and submitting test scores, or they have limited information on college quality and financial aid opportunities (Hoxby and Avery, 2012). The tradeoff analysis with heterogeneous test-taking can be easily modified to address these factors as they impact applicants with varying levels of public endowments in a similar manner.

6.4 Extension #3: An Extra Dimension in Application Profile

One simplification we make in the college application problem is that an applicant's application profile only contains two components: a private endowment (standardized test) and a public endowment (non-test attributes). Then, in our analysis, we use the public endowment as a proxy of socioeconomic status and the private endowment as a proxy of academic merit. In practice, other information required in an application (e.g., high school GPA) may also signal an applicant's merit. Thus, we extend our analysis to incorporate an extra dimension of information available to the school.

Theoretical Argument. Suppose each student's application profile has three components: a private endowment A, two public endowments B_1 and B_2. A represents standardized test scores, B_1 represents public information that signals academic merit (e.g., high school GPA), and B_2 represents public information that signals socioeconomic status (e.g., family zip code). Similar to before, we assume that, all else equal, the school prefers an admission cohort with better academic merit and a more diverse cohort. Let an applicant's true academic merit be M. We assume that the school can perfectly identify M_i if it observes A_i, $B_{1,i}$, $B_{2,i}$ of applicant i, but cannot do so if it only observes $B_{1,i}$ and $B_{2,i}$. It is fair to make this assumption as previous literature has established that standardized test scores help predict students' academic performance in college as well as career success after controlling for other observable student attributes (Bettinger, Evans and Pope, 2013; Chetty, Deming and Friedman, 2023; Cascio et al., 2024; Friedman, Sacerdote and Tine, 2024).

Assuming the school admits the top half applicants with the highest $M + B_2$, Figure

12 shows the composition of the admission cohort under different test policies. Under test-required, the school observes the true merit of every prospective student, so it can admit students in the blue shaded area and achieve the first best (as shown in Panel (a) of Figure 12. Under test blind, the school does not observe A and has to form a belief of merit based on B_1 and B_2. Panel (b) illustrates the difference in admission outcome between test-required (perfect information) and test-blind or test-optional (imperfect information). The hollow circles are examples of applicants that would have appeared high in true student merit but end up low in $E[M|B_1, B_2]$ because they have high A but low B_1. The solid squares are examples of those that would have appeared low in true student merit but end up high in $E[M|B_1, B_2]$ because they have low A but high B_1. If the school has perfect information about M, it should accept most applicants represented by solid and hollow circles and reject most represented by solid squares, as shown in the blue shaded area. However, if the school does not perfectly identify student merit, it may perceive the solid squares high in merit and hollow circles low in merit. As a result, it would accept too many applicants with high B_1 and low M, and reject too many applicants with high M. More specifically, for the solid circles that have high enough B_2, they would still be admitted while their M is underestimated, but for those hollow circles without high enough B_2, they may be rejected by the admission office even if they fall within the shaded area and would have been admitted when school has perfect information. Similarly, solid squares with very low B_2 would not be admitted while their M is overestimated, but for solid squares with not-too-low B_2, they may be admitted even if they fall outside the shaded area because their M is overestimated. In short, the test-blind or test-optional school replaces the hollow circles in the blue shaded area with the solid squares outside the blue shaded area. It is obvious from the graph that every such hollow circle has a higher realized $U = M + B_2$ than every such solid square, and thus the school will have a strict net loss in the realized admission outcomes.

Simulation. To supplement our theoretical argument, we simulate admission outcomes under different test policies when an application profile has three components. Let each student have real quality M in academic preparedness, while the application profile displays private endowment A, public endowments B_1 signaling academic merit, and B_2 signaling socioeconomic status. For simplicity, we assume B_1 and B_2 are independent of each other. M is an integer uniformly distributed between 1 and 5. Suppose with p_b probability $B_1 = M$, with $(1 - p_b)$ probability B_1 is a random integer from the set $\{1, 2, 3, 4, 5\}$; also, with p_a probability $A = M$, with $(1 - p_a)$ probability A is a random integer from the set $\{1, 2, 3, 4, 5\}$. The school admits top half applicants based on $B_2 + E(M|B_1, A)$. Using the Bayes's rule, we can derive the following conditional belief for test-blind and test-required:[12]:

$$\begin{aligned}
&\text{under TB,} \quad E(M|B_1) = 3 - 3p_b + p_b B_1 \\
&\text{under TR, if } A = B_1, \\
&\qquad E(M|B_1, A) = \frac{B_1 \cdot \frac{1}{5}(1+4p_a)(1+4p_b) + (15 - B_1) \cdot \frac{1}{5}(1-p_a)(1-p_b)}{1+4p_a p_b} \\
&\text{under TR, if } A \neq B_1, \\
&\qquad E(M|B_1, A) = \frac{3(1-p_a)(1-p_b) + p_a A + p_b B_1 - p_a p_b (A + B_1)}{1 - p_a p_b}
\end{aligned} \tag{16}$$

Then, we run application and admission simulations under test-blind, test-required, and test-optional policies using a procedure similar to those in Section 6.1. Applicants make reporting decision based on results in Table 4, Model 3.[13] To be consistent with our experimental results, under test-optional, we assume the school belief as follows: if A is reported, the school assigns total endowment as $B_2 + E(M|B_1, A)$; if A is not reported, with probability α, the school assigns total endowment as $B_2 + E(M|B_1)$; with probability $(1-\alpha)$, the school assigns $B_2 + \gamma_0 + \gamma_1 x_{-1} + \gamma_2 y_{-1}$.

Figure 13 presents simulation results when we set $p_b = 0.5$ and $p_a \in \{0.2, 0.5, 0.8\}$. In other words, with 50% probability $B_1 = M$, with 50% probability B_1 is a random integer from the set $\{1, 2, 3, 4, 5\}$. Panels (a)-(c) shows the academic preparedness and diversity tradeoff when the standardized test scores are better ($p_a = 0.8$), the same ($p_a = 0.5$), or worse ($p_a = 0.2$) proxies than high school GPAs ($p_b = 0.5$) of the true academic merit. Similar to what we have found in the simulations with two components in each application profile (Figure 7), Panels (a) and (b) of Figure 13 show that a test-required policy leads to an admission cohort with the highest average academic merit, as well as one of the most diverse cohorts in non-test attributes. Test-blind, however, is strictly dominated by test-required and some test-optional policies. Also similar to Figure 7, test-optional policies that impose the most severe punishment on non-reporting (e.g., C3, C8) admit students with the most diverse non-test attributes. As we have discussed earlier, when subjects are not aware of the harsh interpretation, they do not report their standardized test scores as much as what's best for them. This benefits "low-income high-achieving" applicants while those with higher non-test attributes get severely punished for not reporting. Thus, when the standardized test scores are decent or good proxies of true academic merit, as

[12] Details of the derivation are provided in the appendix.

[13] We replace B in Table 4, Model 3 by B_1. Thus, we assume subjects make reporting decision based on B_1, A, and individual characteristics just as they did based on B, A, and individual characteristics. In other words, the subject's reporting decision is independent from B_2.

found in a recent Dartmouth study (Sacerdote, Staiger and Tine, 2024), our main results are robust to the introduction of a third dimension in the college application profile.

Nevertheless, when standardized tests are less correlated with true academic merit ($p_a = 0.2$), we observe similar admission cohorts under test-blind and test-required in Panel (c). The extra information does not help schools to better identify students with better academic preparedness because it is a very noisy proxy. While test-required no longer leads to admission cohort with the highest academic merit, it also produces a less diverse cohort than many test-optional policies. Therefore, for schools such as the University of California system who have argued that standardized tests do not provide valuable information about applicants' true academic merit, it might be worthwhile to explore other options, such as switching to test-blind or designing new standardized tests that better predict student's academic success.

Figure 14 shows the simulated tradeoff when applicants have perfect information on the school's belief and act rationally accordingly. Similar to what we observe above, test-blind is dominated by test-required and test-optional policies regardless of how good standardized tests predict true merit. When standardized tests are very informative of true merit ($p_a = 0.8$), test-required admits both one of the best academically prepared and one of the most diverse cohorts. [14] When they are as good as high school GPAs in informing merit ($p_a = 0.5$), we observe tradeoff between test-required and some test-optional policies. Nevertheless, the admission cohort under test-required remains both diverse and academically prepared. These suggest that the desirability of test-optional policies shown in Figure 13 mostly comes from applicants not knowing the true school belief. However, when standardized tests are a very noisy proxy of merit ($p_a = 0.2$), the tradeoff in Figure 13 does not change. Many test-optional policies admit a more diverse cohort than test-required, usually at the expense of academic merit.

7 Conclusion

In the past few years, many universities have dropped their SAT or ACT requirements, switching to a test-optional or test-blind admission procedure. While schools have different objective functions with respect to the composition of their entering cohorts, it is unclear how a school should select an admission policy that best represents its interests.

[14] In our simulations, test-required is similar to some test-optional policies (C3, C8) in the academic preparedness and diversity of its admission cohort. Due to the small simulation sample (i.e., 16 applicants per application round), we could observe either C3 and C8 marginally dominate test-required or the other way around.

A major decision for test-optional schools to make is: how to interpret non-reported standardized tests?

In this paper, we study students' reporting choices given their application package and the school's admission statistics from the past, and how these choices drive the school's final admission outcomes. More importantly, we study how student reporting and admission outcomes may differ when the school commits to different interpretations of non-reporting. To overcome the endogeneity criticisms that may come from using observational data, we run a series of controlled lab experiments in a large public university. We address a single-college application problem, in which a student subject's application package has two components: standardized test score (private endowment) and non-test attribute (public endowment), and the school admits students with the highest perceived sum of these two endowments.

We find that the voluntary disclosure of standardized test scores is far from complete, and that this is because the school (our back-end computer) does not give sufficient punishment to non-reporting. Although our experiments do not disclose to student subjects how the school would interpret non-reporting, they manage to learn about the hidden rule. The extent to which they withhold the private endowment is dependent on the hidden school interpretation. Subjects are also more likely to hide their (low) private test scores when they receive a better draw on their observable attribute. Thus, we find evidence of both partial unraveling and reverse unraveling incentives.

Then, we construct a structural model of applicant reporting choice that captures subject learning during the experiment. The model also allows for the possibility that subjects naively follow some rule of thumb and do not engage in strategic learning and reporting from round to round. Using our structural model, we simulate applicants' reporting behavior and admission outcomes under counterfactual school policies. We then discuss a school's tradeoff between admitting students with better academic preparedness and admitting a more diverse cohort. Our simulation suggests that a test-blind policy is dominated by either test-optional or test-required policy in the academic preparedness and the diversity of the admission cohort. Test-required admits students with the highest academic preparedness and from a diverse background. While most test-optional policies are dominated by test-required, some may be desirable when school prioritize the diversity of its entering cohort and severely punishes SAT/ACT non-reporting. However, this gain of diversity (at some cost of academic preparedness) is likely transitory as we only find it present when students are not fully aware of the school's admission policy.

Because the reporting decision is strategic and schools intentionally interpret non-reporting away from its true meaning (in terms of real scores), it encourages lower scores to hide. The more generous the school is in this interpretation, the more it depresses score reporting. On the surface, this may help low-SES students because they have lower scores than high-SES students and would have more incentives to hide if both groups face the same disclosure threshold. But given the positive correlation between family background and test score, high-SES students would have a higher disclosure threshold and therefore are more likely to hide given the same score. This means a low-SES student with a test score may compete against a high-SES student without a score and still lose. In other words, the naive "help" that the test-blind and test-optional policies intend to provide to low-SES students may end up hurting them.

Finally, we extend the stylized college application problem to explore the tradeoff between academic preparedness and diversity under three key modifications: applicant-group admission quotas, resource constraints in test-taking, and an additional dimension in the student application profile. In most cases, our previous findings remain robust: test required is preferable in both dimensions and test blind is the worst, especially if students have perfect information about the school's interpretation of non-reporting. The only exception arises when schools have access to alternative signals of academic ability and standardized test scores are a very noisy measure of true merit. In such cases, test-optional policies can introduce a meaningful tradeoff between average preparedness and diversity as compared to test-required.

In our setting, we assume that all applicants have taken the standardized test and have received their test scores. While other studies have proposed models that incorporate an applicant's endogenous test-taking and school application decisions (Borghesan, 2022), we do not capture student behavior in those margins. The school objective function may also include factors such as social pressure (Dessein, Frankel and Kartik, 2023). We leave the exploration of improved general equilibrium results to future research.

References

Bederson, Benjamin B., Ginger Zhe Jin, Phillip Leslie, Alexander J. Quinn, and Ben Zou. 2018. "Incomplete Disclosure: Evidence of Signaling and Countersignaling." American Economic Journal: Microeconomics, 10(1): 41–66.

Belasco, Andrew S, Kelly O Rosinger, and James C Hearn. 2015. "The test-optional movement at America's selective liberal arts colleges: A boon for equity or something else?" Educational Evaluation and Policy Analysis, 37(2): 206–223.

Bennett, 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.

Bettinger, Eric P., Brent J. Evans, and Devin G. Pope. 2013. "Improving college performance and retention the easy way: Unpacking the ACT exam." American Economic Journal: Economic Policy, 5(2): 26–52.

Board, Oliver. 2009. "Competition and disclosure." The Journal of Industrial Economics, 57(1): 197–213.

Borghesan, Emilio. 2022. "The Heterogeneous Effects of Changing SAT Requirements in Admissions: An Equilibrium Evaluation."

Brand, Jennie E, and Yu Xie. 2010. "Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education." American Sociological Review, 75(2): 273–302.

Cascio, Elizabeth, Bruce Sacerdote, Doug Staiger, and Michele Tine. 2024. "Report From Working Group on the Role of Standardized Test Scores in Undergraduate Admissions." Dartmouth University.

Chen, Daniel L, Martin Schonger, and Chris Wickens. 2016. "oTree—An open-source platform for laboratory, online, and field experiments." Journal of Behavioral and Experimental Finance, 9: 88–97.

Chetty, Raj, David J Deming, and John N Friedman. 2023. "Diversifying society's leaders? The causal effects of admission to highly selective private colleges." National Bureau of Economic Research.

Dale, Stacy Berg, and Alan B Krueger. 2002. "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." The Quarterly Journal of Economics, 117(4): 1491–1527.

Dessein, Wouter, Alex Frankel, and Navin Kartik. 2023. "Test-Optional Admissions." arXiv preprint arXiv:2304.07551.

Dranove, David, and Ginger Zhe Jin. 2010. "Quality disclosure and certification: Theory and practice." Journal of Economic Literature, 48(4): 935–963.

Dynarski, Susan, Aizat Nurshatayeva, Lindsay C Page, and Judith Scott-Clayton. 2023. "Addressing nonfinancial barriers to college access and success: Evidence and policy implications." In Handbook of the Economics of Education. Vol. 6, 319–403. Elsevier.

Eyster, Erik, and Matthew Rabin. 2005. "Cursed equilibrium." Econometrica, 73(5): 1623–1672.

Feltovich, Nick, Richmond Harbaugh, and Ted To. 2002. "Too cool for school? Signalling and countersignalling." RAND Journal of Economics, 630–649.

Freeman, Mark, Preston Magouirk, and Trent Kajikawa. 2021. "Applying to college in a test-optional admissions landscape: trends from Common App data." commonapp.org.

Friedman, John, Bruce Sacerdote, and Michele Tine. 2024. "Standardized Test Scores and Academic Performance at Ivy-Plus Colleges."

Garg, Nikhil, Hannah Li, and Faidra Monachou. 2020. "Dropping standardized testing for admissions trades off information and access." arXiv preprint arXiv:2010.04396.

Grossman, Sanford J. 1981. "The informational role of warranties and private disclosure about product quality." The Journal of Law and Economics, 24(3): 461–483.

Hirshleifer, David, and Siew Hong Teoh. 2003. "Limited attention, information disclosure, and financial reporting." Journal of Accounting and Economics, 36(1-3): 337–386.

Holt, Charles A, and Susan K Laury. 2002. "Risk aversion and incentive effects." American Economic Review, 92(5): 1644–1655.

Hoxby, Caroline M, and Christopher Avery. 2012. "The missing" one-offs": The hidden supply of high-achieving, low income students." National Bureau of Economic Research.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin. 2021. "Is no news (perceived as) bad news? An experimental investigation of information disclosure." American Economic Journal: Microeconomics, 13(2): 141–173.

Liang, Annie, Jay Lu, and Xiaosheng Mu. 2021. "Algorithm Design: A Fairness-Accuracy Frontier." arXiv preprint arXiv:2112.09975.

McManus, Brian, Jessica Howell, and Michael Hurwitz. 2023. "Strategic Disclosure of Test Scores: Evidence from US College Admissions. EdWorkingPaper No. 23-843." Annenberg Institute for School Reform at Brown University.

Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." The Bell Journal of Economics, 380–391.

Moreira, Diana, and Santiago Pérez. 2022. "Who Benefits from Meritocracy?" National Bureau of Economic Research.

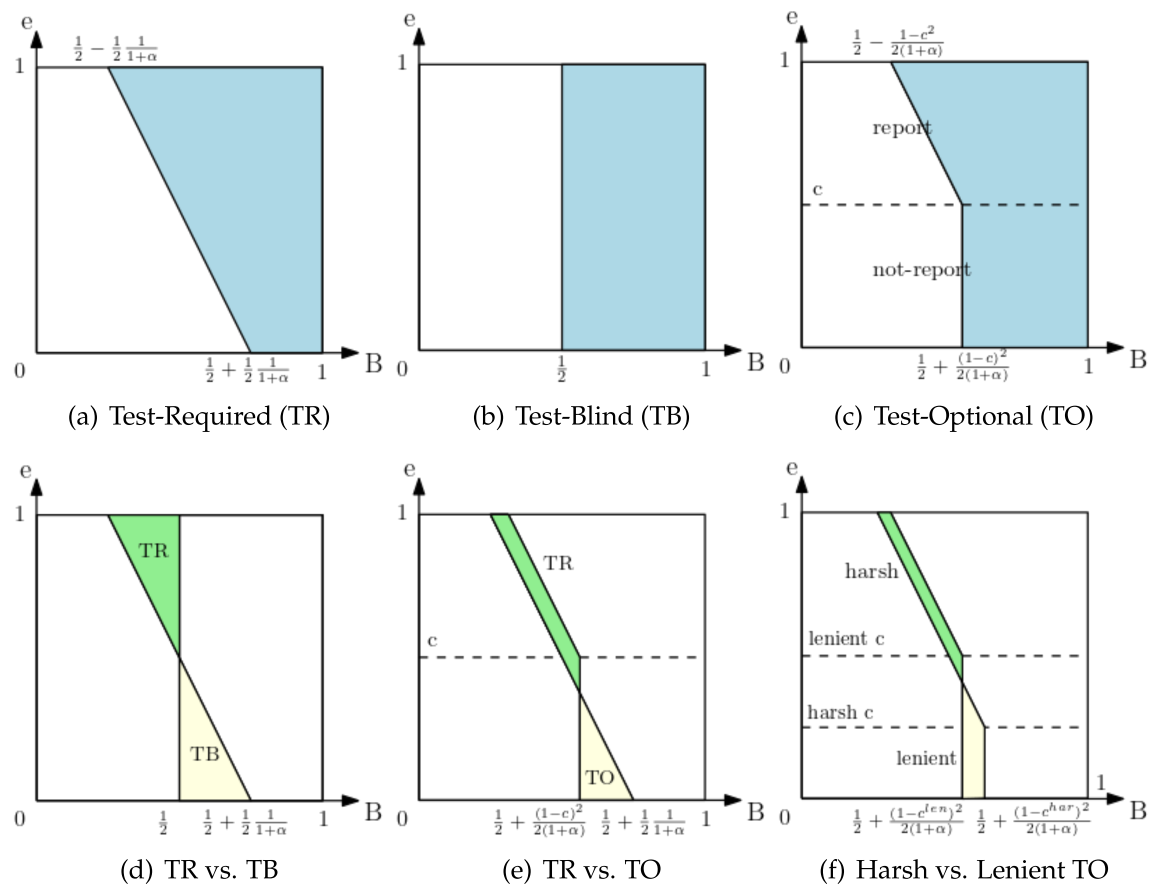
Saboe, Matt, and Sabrina Terrizzi. 2019. "SAT optional policies: Do they influence graduate quality, selectivity or diversity?" Economics Letters, 174: 13–17.

Sacerdote, Bruce, Doug Staiger, and Michele Tine. 2024. "How Test Optional Policies in College Admissions Disproportionately Harm High Achieving Applicants from Disadvantaged Backgrounds."

Sanchez, Henry, and Eddie Comeaux. 2020. "Report of the UC Academic Council Standardized Testing Task Force." University of California Systemwide Academic Senate.

Westrick, Paul A, Jessica P Marini, Doron Shmueli, Linda Young, Emily J Shaw, and Helen Ng. 2020. "Validity of SAT® for Predicting First-Semester, Domain-Specific Grades." College Board.

Figures and Tables

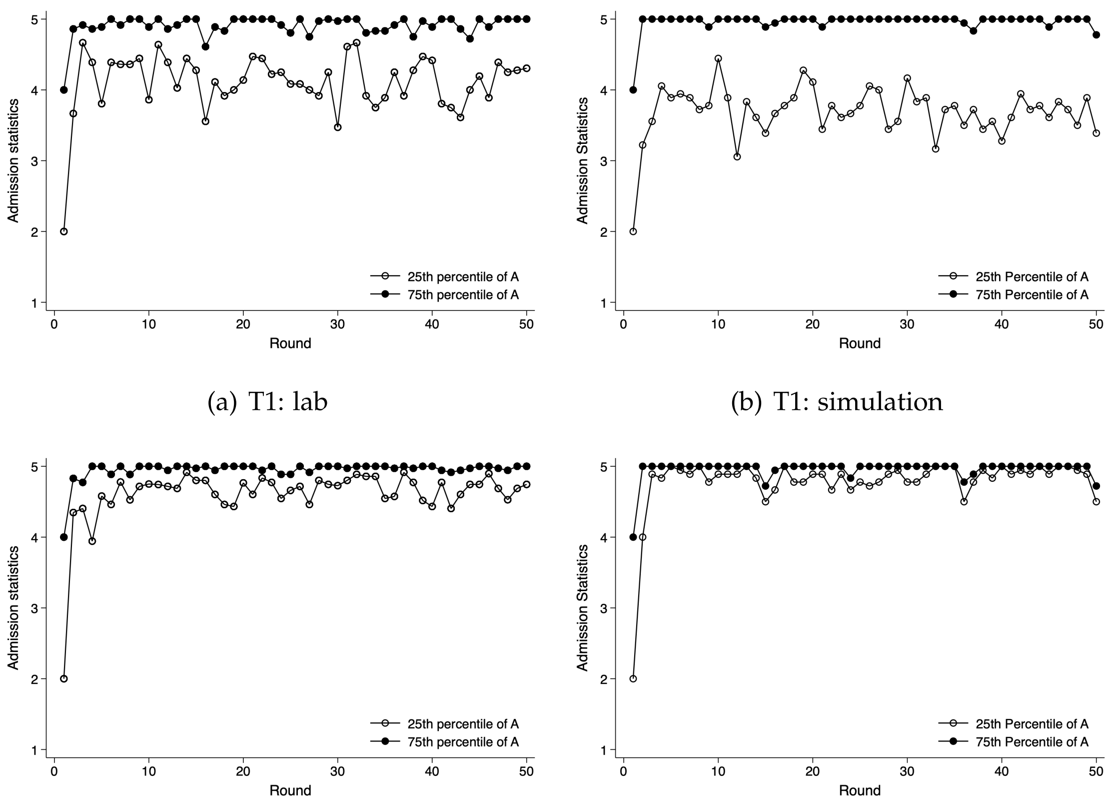


(a) Test-Required (TR) (b) Test-Blind (TB) (c) Test-Optional (TO)

(d) TR vs. TB (e) TR vs. TO (f) Harsh vs. Lenient TO

Figure 1: Composition of the Admission Cohort

Note: This figure provides an illustration of the composition of the admission cohort under test-required, test-blind, and test-optional admission policies. Assuming the school admits half of all applicants, the blue shaded areas in Panels (a), (b), and (c) represent admitted applicants. In Panels (c), (d), and (e), we show the difference in admission composition between test policy pairs.



(a) T1: lab

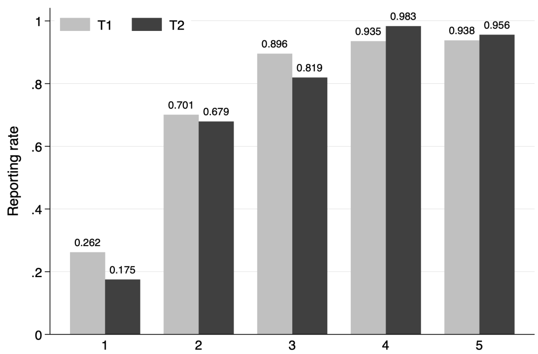
(b) T1: simulation

(c) T2: lab

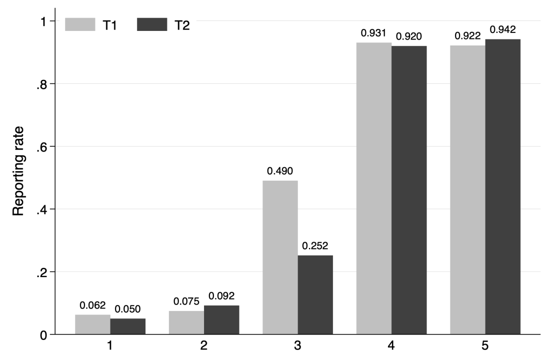
(d) T2: simulation

Figure 2: Trends in Admission Statistics

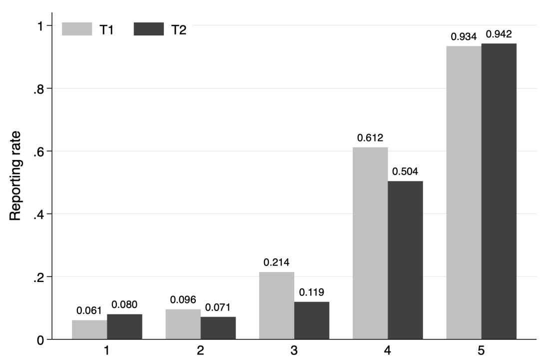
Note: This figure shows the trends in admission statistics by round. Panels (a) and (c) present the trends in the 25th and 75th percentiles of private endowment from those who reported it and were offered admission in the lab under T1 and T2, respectively. Panels (b) and (c) present the trends when we simulate subject reporting decision with perfect information under T1 and T2, respectively.



(a) B=1



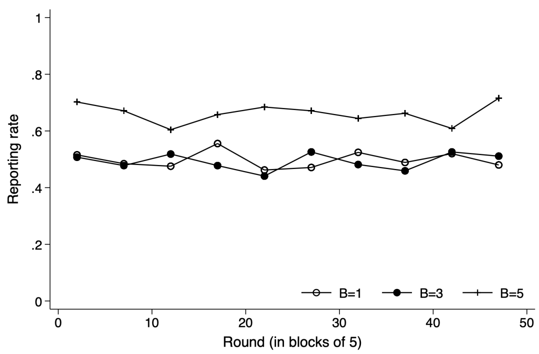
(b) B=3



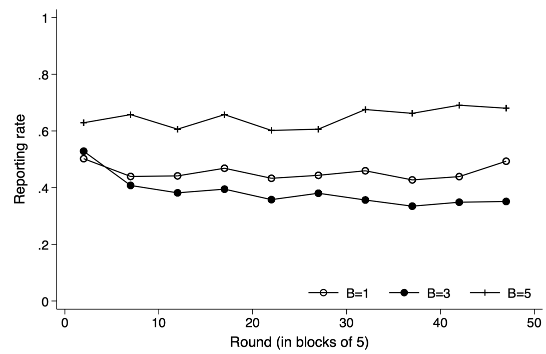
(c) B=5

Figure 3: Reporting Rates in Rounds 21-50

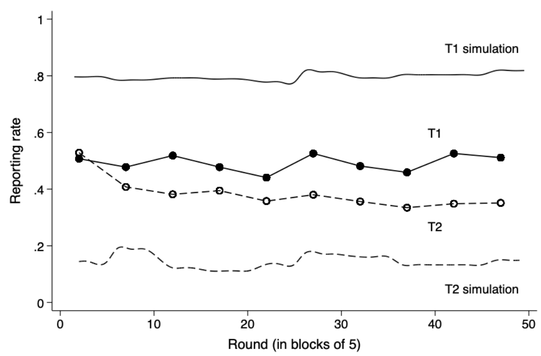
Note: This figure presents the reporting rates by public and private endowments in the last 30 rounds of the experiment. Each panel represents one possible public endowment (B) and the horizontal axis in each panel represents the private endowment (A). The light gray bars show the reporting rates under T1, and the dark gray bars show the reporting rates under T2.



(a) T1: lab



(b) T2: lab



(c) B=3

Figure 4: Trends in Subject Reporting Rates

Note: This figure shows the trends in subject reporting rates in the lab. We present a separate trend for each possible public endowment (B) under T1 and T2. The reporting rates are averaged across every 5 rounds.

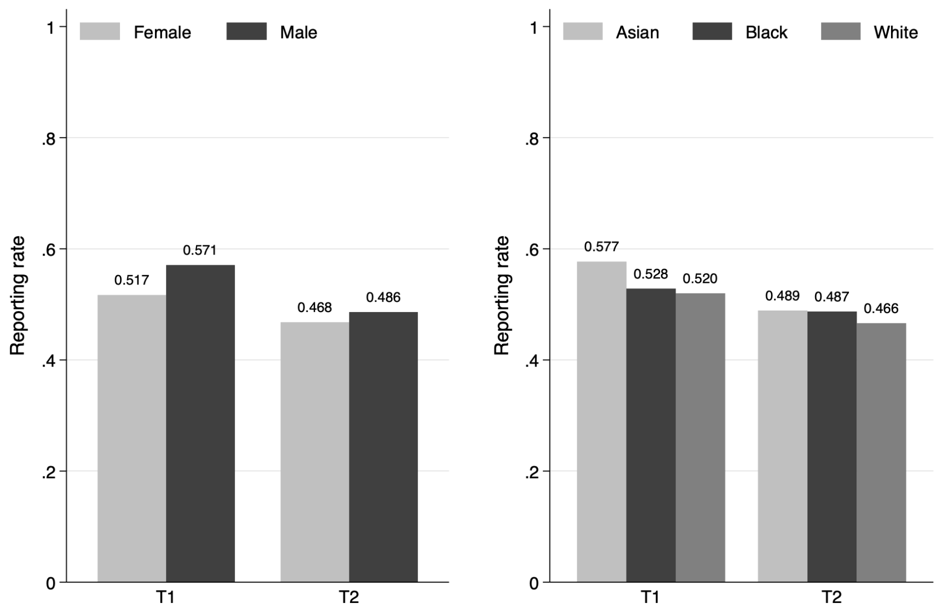
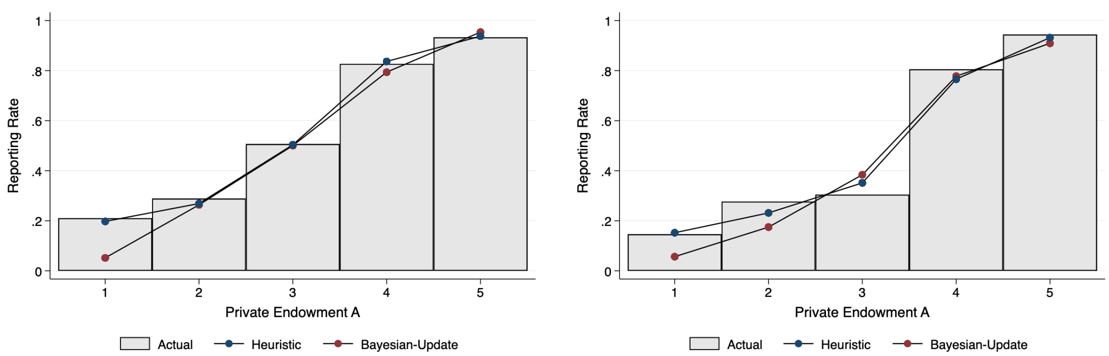


Figure 5: Reporting Rates by Gender and Race

Note: This figure presents the reporting rates under T1 and T2 by (self-reported) gender and race in the last 30 rounds of the experiment. The graph on the left shows the reporting rates for female and male. The graph on the right shows the reporting rates for Asian, Black, and White subjects.



(a) T1

(b) T2

Figure 6: Model Predicted Reporting Rates

Note: This figure shows the actual and the predicted reporting rates by the heuristic model and the Bayesian updating model.

Figure 7: Academic Preparedness and Diversity Tradeoff

Note: This figure shows the academic preparedness and diversity of the admission cohorts under various admission policies, based on simulation results from our structural model. The "academic preparedness" is defined as the average of the private endowment, and "diversity" is defined as the standard deviation of the public endowment. Each point on the graph represents one possible admission policy. The admission policies include two treatments in our experiment (T1, T2), test-required (TR), test-blind (TB), and the counterfactual test-optional policies (C1-C16).

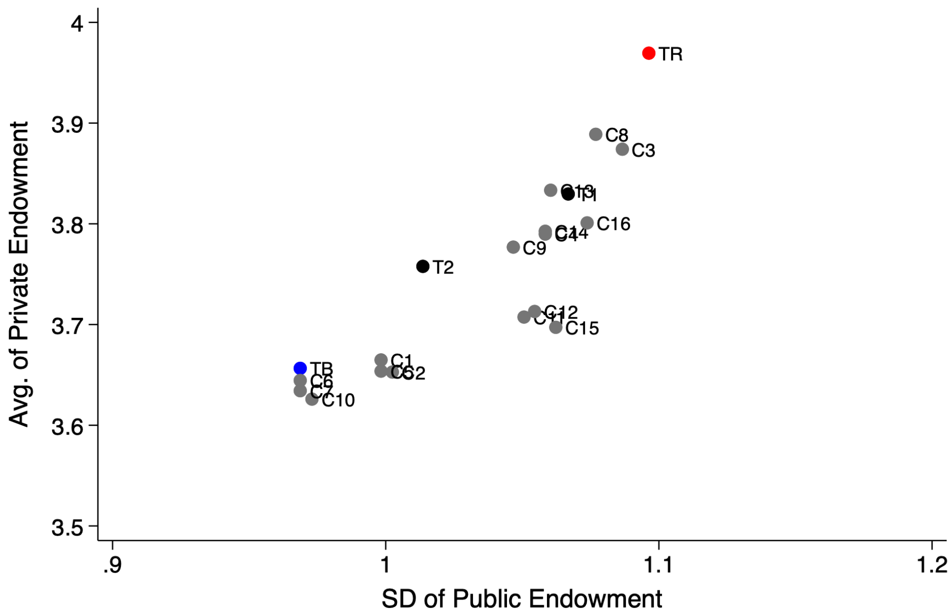
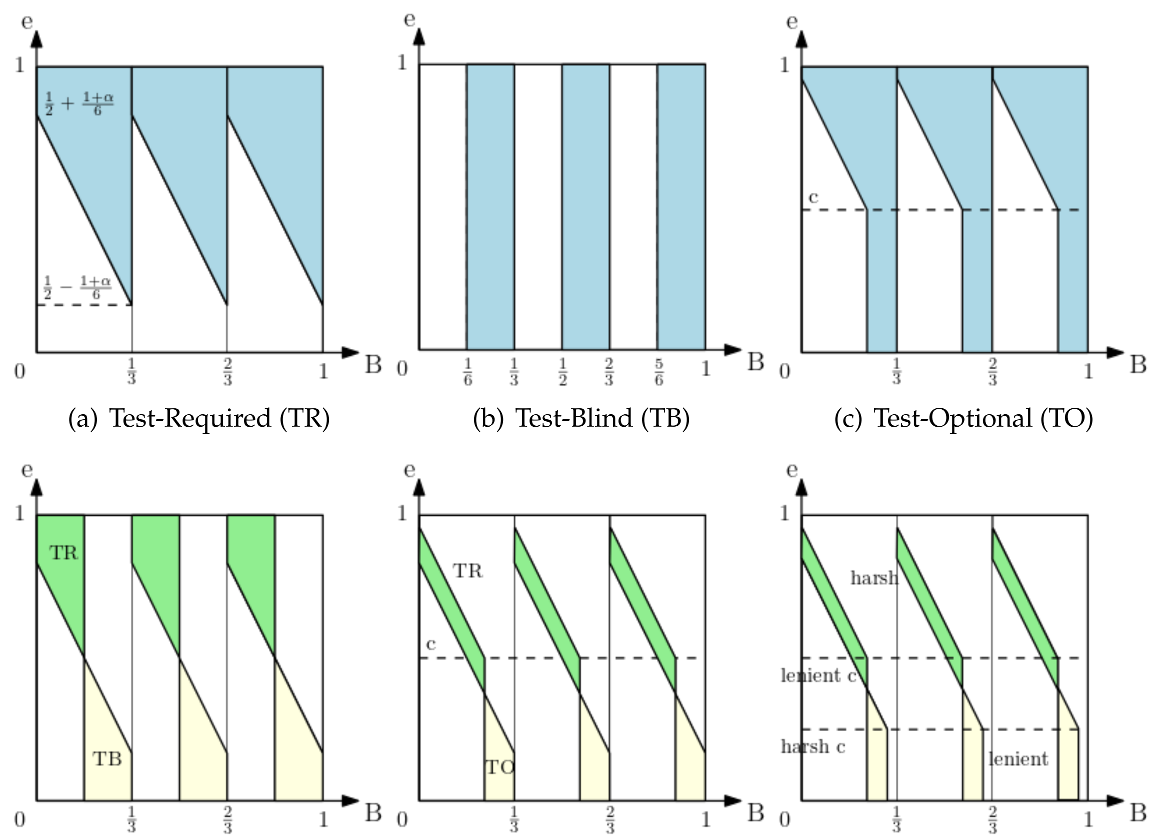


Figure 8: The Tradeoff with Perfect Information

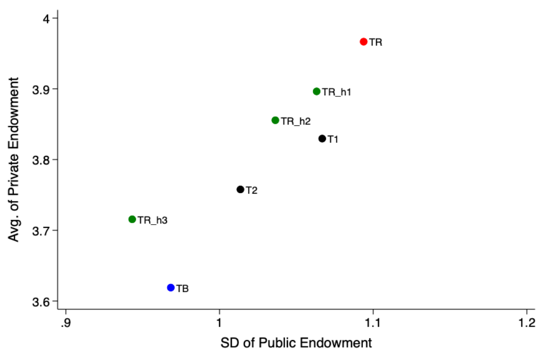
Note: This figure shows the academic preparedness and diversity of the admission cohorts under various admission policies, assuming that applicants have perfect information on the school's interpretation of non-reported standardized tests.



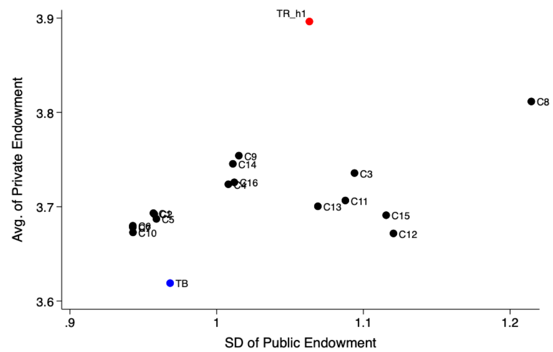
(a) Test-Required (TR)
(b) Test-Blind (TB)
(c) Test-Optional (TO)
(d) TR vs. TB
(e) TR vs. TO
(f) Harsh vs. Lenient TO

Figure 9: Composition of the Admission Cohort

Note: This figure provides an illustration of the composition of the admission cohort under test-required, test-blind, and test-optional admission policies when there is an admission quota on the public endowment. In particular, the school admits one third of its cohort from each of low, medium, and high public endowment. Assuming the school admits half of all applicants, the blue shaded areas in Panels (a), (b), and (c) represent admitted applicants. In Panels (c), (d), and (e), we show the difference in admission composition between test policy pairs.

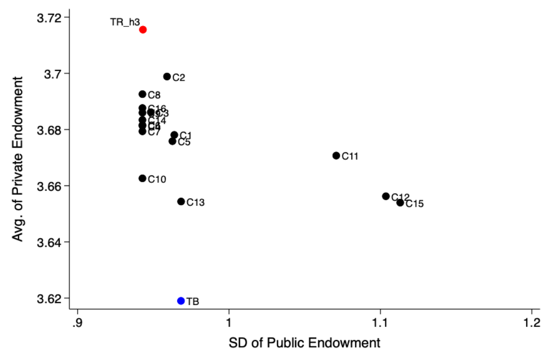


(a) TR under Heterogeneous Test-Taking



(b) $p_1 = 0.5$, $p_3 = 0.75$, $p_5 = 1$

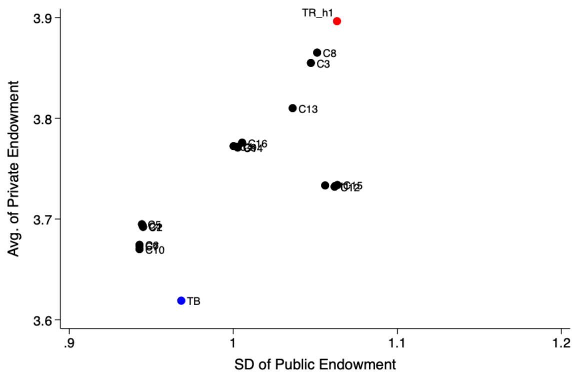
(c) $p_1 = 0.33$, $p_3 = 0.67$, $p_5 = 1$



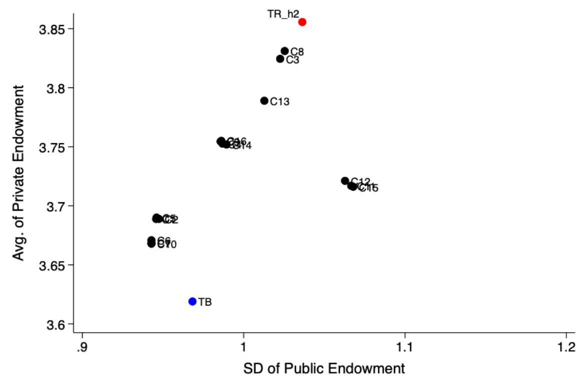
(d) $p_1 = 0$, $p_3 = 0.5$, $p_5 = 1$

Figure 10: Tradeoff with Heterogeneous Test-Taking

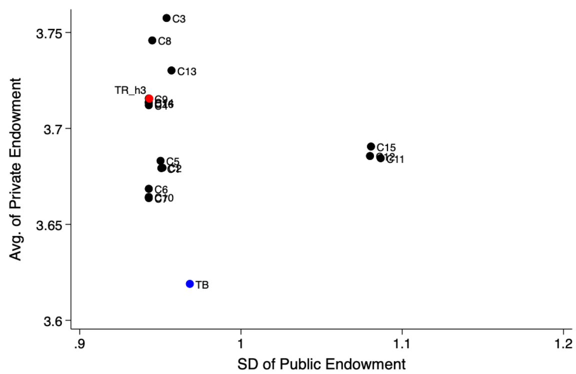
Note: This figure presents the academic preparedness and diversity tradeoff with heterogeneous test-taking behavior. In Panel (a) we show how test-required compares to test-optional and test-blind with different sets of test-taking probabilities. In Panels (b), (c), and (d) we use our model to simulate the reporting behavior and admission results when a proportion of applicants does not take standardized tests.



(a) $p_1 = 0.5$, $p_3 = 0.75$, $p_5 = 1$



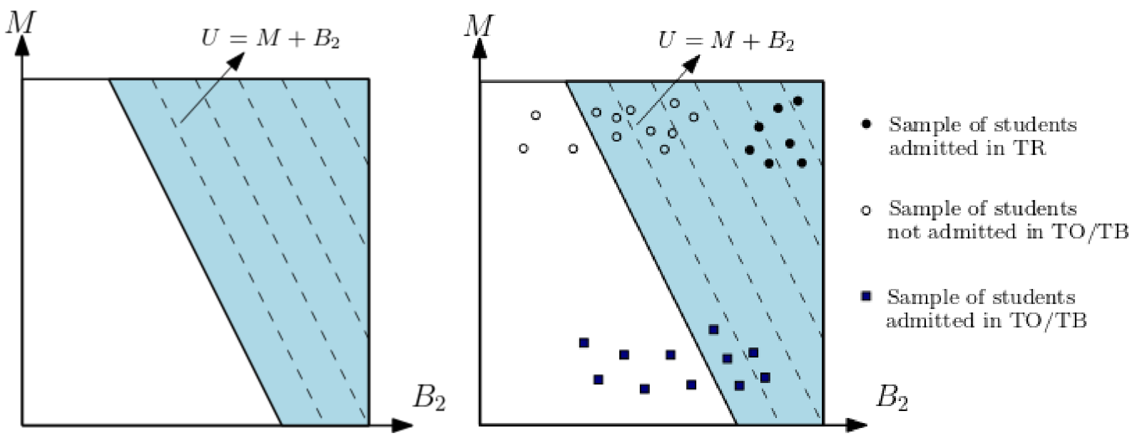
(b) $p_1 = 0.33$, $p_3 = 0.67$, $p_5 = 1$



(c) $p_1 = 0$, $p_3 = 0.5$, $p_5 = 1$

Figure 11: Tradeoff with Heterogeneous Test-Taking and Perfect Information

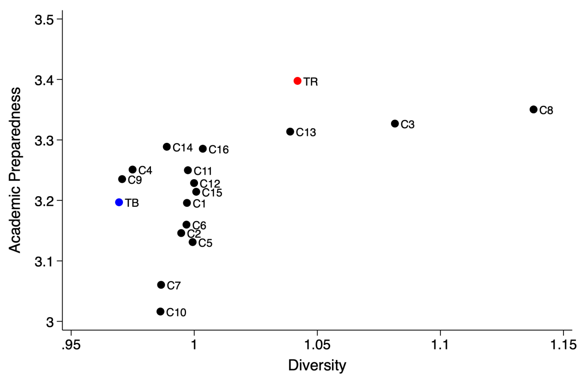
Note: This figure presents the academic preparedness and diversity tradeoff with heterogeneous test-taking behavioral when subjects have perfect information on the school's interpretation and act rationally accordingly.



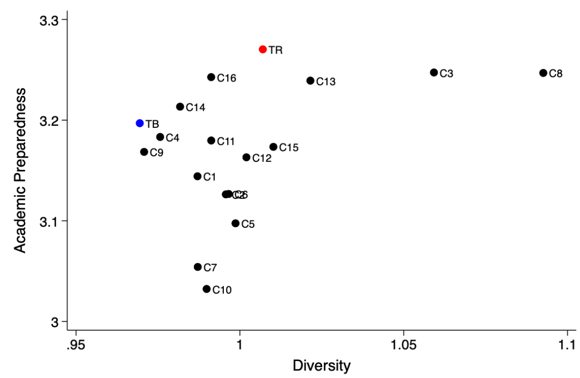
(a) Perfect Info. on Student Merit (b) Imperfect Info. on Student Merit

Figure 12: Composition of the Admission Cohort with an Extra Dimension

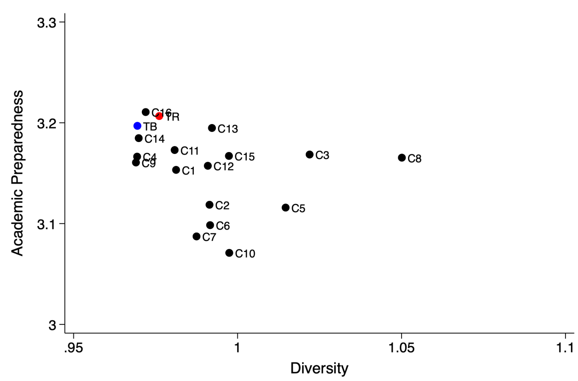
Note: This figure illustrates the composition of the admission cohort when there's three components in an applicant's application profile. Panel (a) shows the composition under a test-required (TR) policy. Panel (b) shows the composition under either test-optional (TO) or test-blind (TB).



(a) $p_a = 0.8$, $p_b = 0.5$



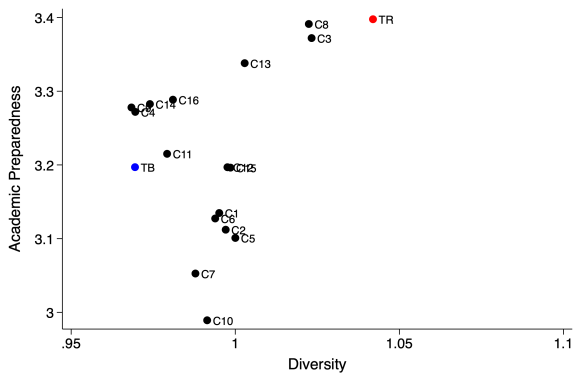
(b) $p_a = 0.5$, $p_b = 0.5$



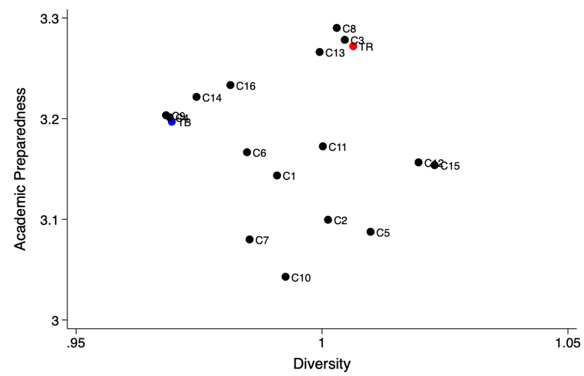
(c) $p_a = 0.2$, $p_b = 0.5$

Figure 13: The Tradeoff with An Extra Dimension

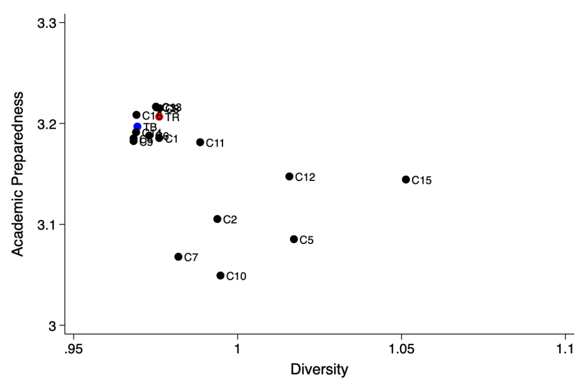
Note: This figure shows the academic preparedness and diversity of the admission cohorts under various admission policies when we incorporate an extra dimension of information available to the school. Panel (a) shows simulation results when $p_a = 0.8$, $p_b = 0.5$. Panels (b) and (c) show simulation results when $p_a = 0.5$, $p_b = 0.5$ and $p_a = 0.2$, $p_b = 0.5$, respectively.



(a) $p_a = 0.8$, $p_b = 0.5$



(b) $p_a = 0.5$, $p_b = 0.5$



(c) $p_a = 0.2$, $p_b = 0.5$

Figure 14: Tradeoff with An Extra Dimension and Perfect Information

Note: This figure shows the academic preparedness and diversity tradeoff when we incorporate an extra dimension of information and when applicants have perfect information on the school's belief and act rationally accordingly.

Table 1: Summary Statistics of Main Sessions

	T1 (1)	T2 (2)	T1-T2 (3)	p-value (4)
Demographics				
Female	0.53	0.56	-0.03	0.58
Asian	0.32	0.32	0.00	1.00
Black	0.14	0.09	0.05	0.22
White	0.49	0.49	-0.00	0.96
Others	0.06	0.10	-0.04	0.17
Freshman	0.38	0.41	-0.04	0.53
Sophomore	0.17	0.19	-0.02	0.70
Junior	0.21	0.23	-0.03	0.60
Senior	0.24	0.16	0.08	0.09
Standardized Test				
Number of Attempts	1.90	1.85	0.04	0.77
SAT Math Score	693.10	684.19	8.91	0.47
SAT Reading Score	677.54	689.29	-11.75	0.25
ACT Score	29.53	31.43	-1.90	0.04
Number of Schools Applied	7.52	7.77	-0.25	0.65
Non-Report as Bad Signal	0.53	0.65	-0.12	0.03
Send SAT/ACT to All Schools	0.47	0.43	0.04	0.50
Did Not Take SAT/ACT	0.12	0.17	-0.06	0.32
School Test-Blind	0.08	0.07	0.00	0.93
Had Low Scores	0.57	0.65	-0.08	0.29
UMD as First Choice	0.33	0.33	0.01	0.90
N	144	141	285	

* *Notes:* This table reports the summary statistics of subjects in our experiments. The first two columns show the summary statistics of demographics and experience with college applications for subjects in treatments T1 and T2, respectively. Column 3 shows the difference between the first two columns, and Column 4 presents the p-value for the difference.

Table 2: Summary of Subject Reporting Rates

Treatment	T1			T2		
	All Rounds	Rounds 1-20	Rounds 21-50	All Rounds	Rounds 1-20	Rounds 21-50
Panel A: avg. reporting rate by private endow.						
A=1	0.217	0.227	0.210	0.156	0.171	0.146
A=2	0.294	0.304	0.288	0.283	0.292	0.277
A=3	0.490	0.465	0.506	0.344	0.405	0.304
A=4	0.818	0.806	0.827	0.793	0.775	0.805
A=5	0.935	0.937	0.933	0.939	0.931	0.944
Panel B: avg. reporting rate by public endow.						
B=1	0.498	0.508	0.491	0.455	0.463	0.449
B=3	0.493	0.495	0.491	0.384	0.428	0.355
B=5	0.662	0.659	0.664	0.647	0.638	0.653
Total N	7200	2880	4320	7050	2820	4230

[*] *Notes:* This table reports a summary of subject reporting rates during the experiment. For each private or public endowment, we report the average reporting rates of private endowment in all rounds, in the first 20 rounds, and in the last 30 rounds.

Table 3: Hypothetical Reporting Rule with Perfect Information

Public Endowment	School-Belief	Report under T1	Report under T2
$B_t = 1$	T1: $0.5 + 0.25x_{t-1}$ ≈ 1.5 T2: $0.5 + 0.25(x_{t-1} + y_{t-1})$ ≈ 2.875	if $A_t \in \{2, 3, 4, 5\}$	if $A_t \in \{3, 4, 5\}$
$B_t = 3$	T1: $1.5 + 0.25x_{t-1}$ ≈ 2.5 T2: $1.5 + 0.25(x_{t-1} + y_{t-1})$ ≈ 3.875	if $A_t \in \{3, 4, 5\}$	if $A_t \in \{4, 5\}$
$B_t = 5$	T1: $2.5 + 0.25x_{t-1}$ ≈ 3.5 T2: $2.5 + 0.25(x_{t-1} + y_{t-1})$ ≈ 4.875	if $A_t \in \{4, 5\}$	if $A_t \in \{5\}$

* *Notes:* This table reports the hypothetical reporting rule (indicated by Equation 10) when subjects have perfect information about the school's interpretation of non-reported private endowment. Column 2 reports the true school-belief under T1 and T2 given the public endowment in Column 1. Columns 3 and 4 show the values of private endowment with which subjects with perfect information would report under T1 and T2, respectively.

Table 4: Logit Model Estimation of Reporting Decisions

	(1) All Rounds	(2) Rounds 21-50	(3) Rounds 21-50	(4) Rounds 21-50
Public Endow. B	-0.34***	-0.34***	-0.72***	-0.69***
	(0.01)	(0.01)	(0.08)	(0.10)
Q1 (previous round)	-0.20***	-0.23***	-0.14***	-0.09
	(0.03)	(0.04)	(0.05)	(0.06)
Q3 (previous round)	-0.10	-0.12	-0.12	-0.19
	(0.08)	(0.11)	(0.14)	(0.15)
Learning			X	X
Subject Covariates			X	
Subject FE				X
N	14250	8550	8550	8550

[*] *Notes:* This table reports estimates from a logistic regression of subject reporting decision. In Columns 3 and 4 we include eight learning variables. The subject covariates in Column 3 include those summarized in Table 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Heuristic Structural Model Predicted Reporting Rates

Panel A: Reporting rate (%)	All				Treatment: T1		Treatment: T2	
	Actual	Model (2)	Model (3)	Model (4)	Actual	Model (3)	Actual	Model (3)
A=1	17.9	19.1	17.5	18.0	21.0	19.7	14.6	15.2
A=2	28.3	29.8	25.1	25.3	28.8	26.9	27.7	23.1
A=3	40.6	51.8	42.8	42.6	50.6	50.4	30.4	35.1
A=4	81.6	72.7	80.2	80.1	82.7	83.7	80.5	76.6
A=5	93.8	82.8	93.5	93.2	93.3	93.8	94.4	93.2
Avg. distance (unweighted)		6.8	1.5	1.4		1.0		3.0
Total log likelihood		-4018	-3029	-2835				

Panel B: (%) with Naivety	All				Treatment: T1		Treatment: T2	
	Actual	Model (2)	Model (3)	Model (4)	Actual	Model (3)	Actual	Model (3)
A=1	17.9	16.4	17.0	17.5	21.0	19.1	14.6	14.8
A=2	28.3	25.6	24.3	24.6	28.8	26.1	27.7	22.4
A=3	40.6	44.6	41.6	41.3	50.6	48.9	30.4	34.1
A=4	81.6	86.6	81.6	81.7	82.7	84.8	80.5	78.2
A=5	93.8	91.6	93.9	93.8	93.3	94.2	94.4	93.7
Avg. distance (unweighted)		3.1	1.2	1.0		1.9		2.4
Naivety (θ_1)		0.51	0.03	0.03				
Naivety (θ_2)		0.14	0.07	0.08				
Total log likelihood		-3864	-3021	-2830				

* *Notes:* This table reports the actual subject reporting rates, the predicted reporting rates from our structural model, and their differences. Panel A reports these statistics from three different model specifications presented in Table 4. Panel B reports the same set of statistics when we allow subjects to be naive with a positive probability. θ_1 represents the probability that subjects never report when their private endowment (A) was less than or equal to 3, regardless of their public endowment (B). θ_2 represents the probability that subjects always report when their private endowment (A) was higher than 3, regardless of their public endowment (B). The naivety parameters are estimated using maximum likelihood estimation.

Table 6: List of School Interpretations of Non-Reporting

Setting	α	γ_0	γ_1	γ_2	$\mathrm{E}[A_t \mid R_t = 0]$
Experiment					
T1	0.5	0	0.5	0	$0.5B_t + 0.25x_{t-1}$
T2	0.5	0	0.5	0.5	$0.5B_t + 0.25(x_{t-1} + y_{t-1})$
Counterfactual					
C1	0.5	0	1	0	$0.5B_t + 0.5x_{t-1}$
C2	0.5	0	0	1	$0.5B_t + 0.5y_{t-1}$
C3	0.5	1	0	0	$0.5B_t + 0.5$
C4	0.5	3	0	0	$0.5B_t + 1.5$
C5	0.5	5	0	0	$0.5B_t + 2.5$
C6	0.25	0	1	0	$0.25B_t + 0.75x_{t-1}$
C7	0.25	0	0	1	$0.25B_t + 0.75y_{t-1}$
C8	0.25	1	0	0	$0.25B_t + 0.75$
C9	0.25	3	0	0	$0.25B_t + 2.25$
C10	0.25	5	0	0	$0.25B_t + 3.75$
C11	0.75	0	1	0	$0.75B_t + 0.25x_{t-1}$
C12	0.75	0	0	1	$0.75B_t + 0.25y_{t-1}$
C13	0.75	1	0	0	$0.75B_t + 0.25$
C14	0.75	3	0	0	$0.75B_t + 0.75$
C15	0.75	5	0	0	$0.75B_t + 1.25$
C16	1	.	.	.	B_t

* *Notes:* This table reports the list of counterfactual school interpretations of non-reported private endowment that we simulated. The variation in these test-optional school policies come from the selection of parameters $\alpha, \gamma_0, \gamma_1, \gamma_2$, which we introduced in Equation 5.

A Appendix: Supplementary Materials

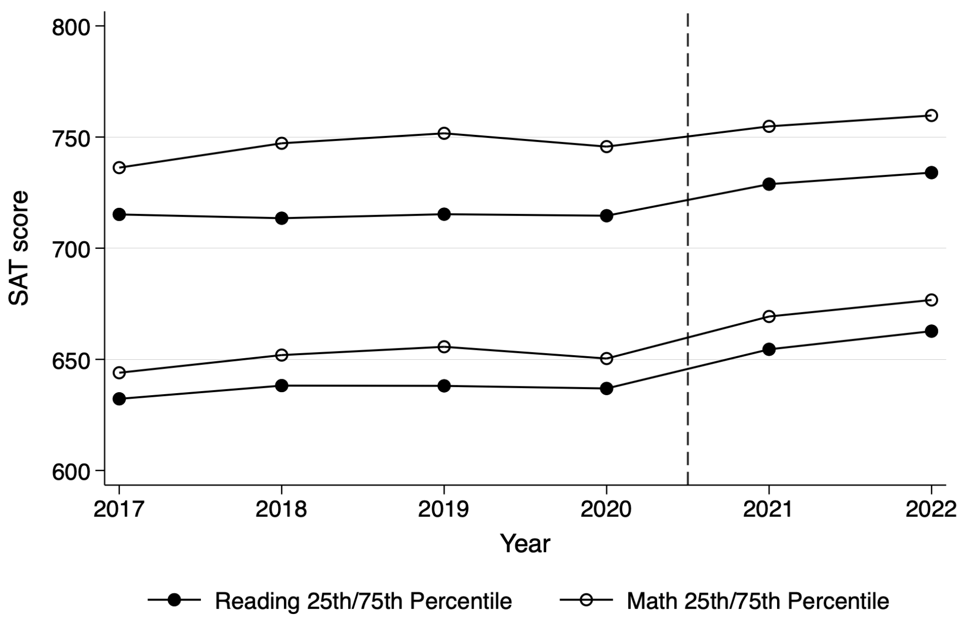


Figure A.1: Trends in Reported SAT Reading and Math Scores

Note: This figure shows the trends in SAT reading and math scores at the top-100 schools ranked by U.S. News in the past five years. The scores in this figure represent the average test scores of those applicants who reported their scores and were admitted to these schools. Source: Department of Education and U.S. News.

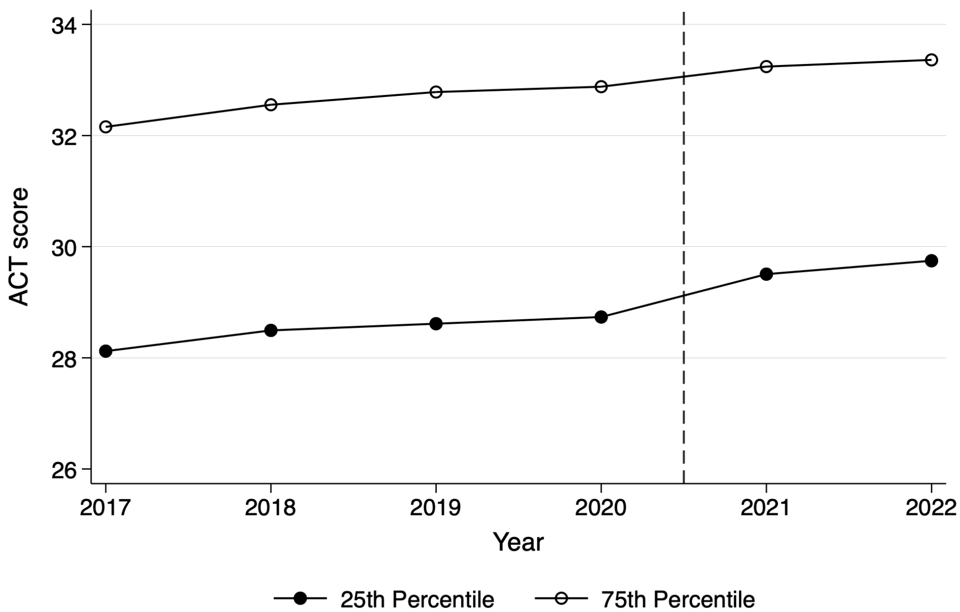


Figure A.2: Trends in Reported ACT Scores

Note: This figure shows the trends in ACT composite score at the top-100 schools ranked by U.S. News in the past five years. The scores in this figure represent the average test scores of those applicants who reported their scores and were admitted to these schools. Source: Department of Education and U.S. News.

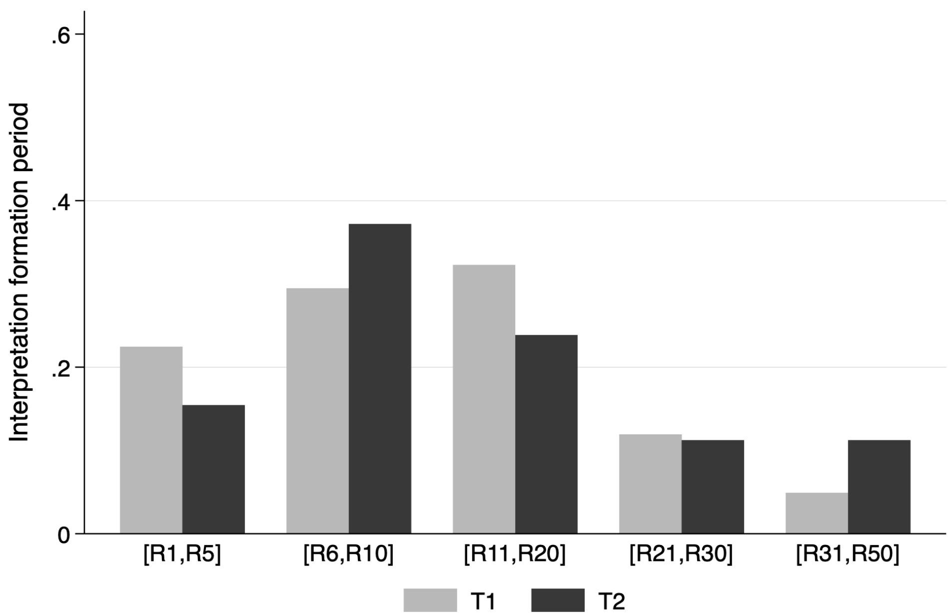


Figure A.3: Time of Belief Formation

Note: This figure shows the distribution of rounds at which subjects (self-reported) formed their beliefs of the school-interpretation of non-reported private endowment. Specifically, the statistics are drawn from the responses to this survey question at the end of the experiment: "When, during the experiment, did you realize the program's interpretation?".

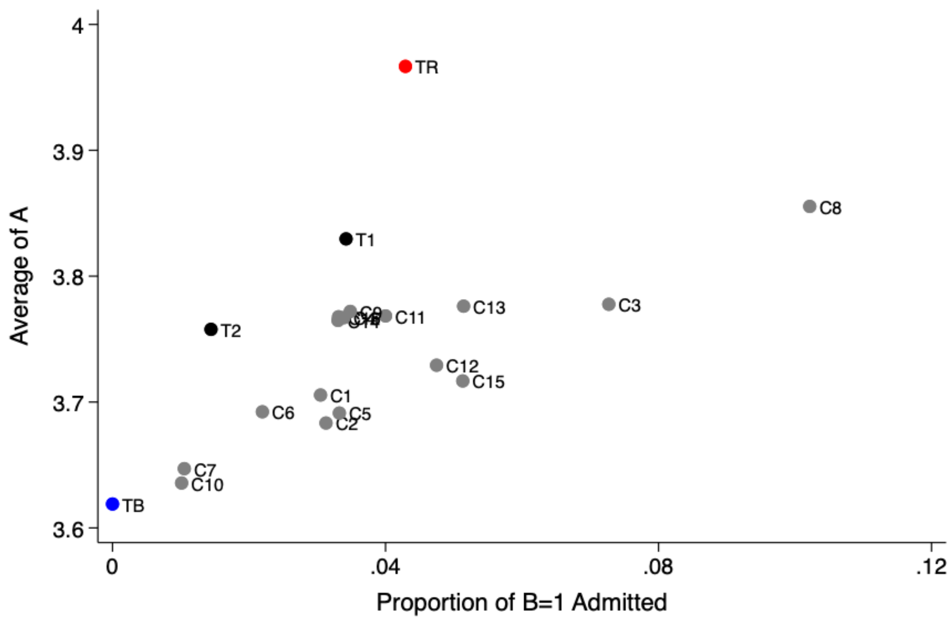


Figure A.4: Academic Preparedness and Diversity Tradeoff (Alternative Diveristy Measure)

Note: This figure shows the academic preparedness and diversity of the admission cohorts under various admission policies, based on simulation results from our structural model. The "academic preparedness" is defined as the average of the private endowment, and "diversity" is defined as the proportion of $B = 1$ subjects admitted. Each point on the graph represents one possible admission policy. The admission policies include two treatments in our experiment (T1, T2), test-required (TR), test-blind (TB), and the counterfactual test-optional policies (C1-C16).

Table A1: Bayesian Updating Model Estimates of the Prior

$\phi^{(1)}_{\alpha}$	$\phi^{(1)}_{\alpha=0}$	$\phi^{(1)}_{\alpha=0.25}$	$\phi^{(1)}_{\alpha=0.5}$	$\phi^{(1)}_{\alpha=0.75}$	$\phi^{(1)}_{\alpha=1}$	
	0.001	0.381	0.522	0.010	0.000	
$\phi^{(1)}_{\gamma_0}$	$\phi^{(1)}_{\gamma_0=0}$	$\phi^{(1)}_{\gamma_0=1}$	$\phi^{(1)}_{\gamma_0=2}$	$\phi^{(1)}_{\gamma_0=3}$	$\phi^{(1)}_{\gamma_0=4}$	
	0.000	0.996	0.003	0.001	0.000	
$\phi^{(1)}_{\gamma_1,\gamma_2}$	$\phi^{(1)}_{\gamma_1=0,\gamma_2=0}$	$\phi^{(1)}_{\gamma_1=0,\gamma_2=0.5}$	$\phi^{(1)}_{\gamma_1=0,\gamma_2=1}$	$\phi^{(1)}_{\gamma_1=0.5,\gamma_2=0}$	$\phi^{(1)}_{\gamma_1=1,\gamma_2=0}$	$\phi^{(1)}_{\gamma_1=0.5,\gamma_2=0.5}$
	0.130	0.266	0.000	0.604	0.000	0.000
Covariates	pc1	pc2	pc3	pc4	pc5	pc6
	-0.024	0.037	0.037	-0.030	0.056	0.057
	pc7	pc8	pc9			
	0.053	0.032	-0.114			

* *Notes:* This table reports estimates of the prior from the Bayesian updating model. For the covariates, we perform a principal component analysis and keep the first nine components, which have eigenvalues larger than 1.

Experimental Instructions

Welcome

You are about to participate in an experiment on decision-making, and you will be paid for your participation in cash, privately at the end of the experiment. What you earn depends partly on your decisions, and partly on chance. Please silence and put away your cellular phones now. The entire session will take place through your computer terminal. Please do not talk or in any way communicate with other participants during the session. We will start with a brief instruction period. During the instruction period, you will be given a description of the main features of the experiment and will be shown how to use the computers. If you have any questions during this period, raise your hand and your question will be answered so everyone can hear.

Instructions

1. The experiment you are participating in consists of 50 rounds. At the end of the final round, the computer will select two random rounds, and you will be paid based on your outcomes in those two rounds (in addition to the $10 show-up fee). Everybody will be paid in private. You are under no obligation to tell others how much you earned.

2. You are an applicant for a competitive program. Every round, the program admits half of the applicants based on its belief of each applicant's endowment. In the case of multiple applicants, with equal perceived endowment, competing for one or more admission seats, the program will randomly pick the competing applicants for admission with equal probability. It is in your best interest to be admitted by the program in every round, since it increases the chance of you being paid if that round is selected at random. For each admission you get in the two randomly selected rounds, you will be paid $6 (with a maximum of $12 in total if you get two admissions) in addition to the $10 show-up fee.

3. In each round, the computer will generate two numbers, A and B. A is drawn from the set {1, 2, 3, 4, 5} and B is drawn from the set {1, 3, 5}. These two numbers, A and B, are the only two components of your endowment, and they will be displayed on your screen. The sum of A and B (i.e., A+B) is your overall endowment. A is private information and is unknown to the program. B is public information and is known to the program. You will choose whether or not to report A, having in mind that the program knows your B. If you choose to report A, you can only report the true number. Otherwise, you do not report and send no message to the program.

4. If you report, the program knows both A and B, and identifies your overall endowment perfectly. If you do not report, the program only knows B. Then, the program will form a belief about your A, and subsequently a belief about your overall endowment summing A and B. The program's belief of A is positively correlated with its knowledge of B. In other words, the higher is your B, the higher is the chance that

the program thinks your A is high. The program will admit half of the applicants with the highest overall endowment based on its knowledge or belief of each applicant's overall endowment. After you choose to report or not report A, the admission result will be displayed on your screen.

5. Before you make your choice, the screen will display the following admission statistics from the last round: "Of all the applicants that were admitted to the program in the last round, including those who reported their A and those who did not report their A, the average value of B is Z. Of all the applicants that reported their A, and were admitted to the program in the last round, the 25th percentile of A was X and the 75th percentile of A was Y. The higher the program thinks your endowment is, the higher is the probability that you will be admitted."

6. During the experiment, an understanding of the concepts of 25th percentile and 75th percentile is crucial. In short, the 25th percentile is the value at which 25% of the endowments lie below that value; the 75th percentile is the value at which 75% of the endowments lie below that value. On the first page of your screen, you will see some examples of lists of numbers and their corresponding 25th and 75th percentiles. Please read the examples carefully before you move on to the next page.

Lab Screenshots

Introduction

Instructions:

The experiment you are participating in consists of 50 rounds. At the end of the final round, the computer will select two random rounds, and you will be paid based on your outcomes in those two rounds (in addition to the $10 show up fee). Everybody will be paid in private. You are under no obligation to tell others how much you earned.

You are an applicant for a competitive program. Every round, the program admits half of the applicants based on its belief of each applicant's endowment. In the case of multiple applicants, with equal perceived endowment, competing for one or more admission seats, the program will randomly pick the competing applicants for admission with equal probability. It is in your best interest to be admitted by the program in every round, since it increases the chance of you being paid if that round is selected at random. For each admission you get in the two randomly selected rounds, you will be paid $6 (with a maximum of $12 in total if you get two admissions) in addition to the $10 show up fee.

During the experiment, an understanding of the concepts of 25th percentile and 75th percentile is crucial. In short, 25th percentile is the value at which 25% of the endowments lie below that value; 75th percentile is the value at which 75% of the endowments lie below that value. Here are some examples of lists of numbers and their 25th and 75th percentiles:

List: {1, 2, 3, 4, 5}: 25th percentile is 2.0, 75th percentile is 4.0

List: {1, 1, 3, 5, 5}: 25th percentile is 1.0, 75th percentile is 5.0

List: {1, 3, 3, 3, 5}: 25th percentile is 3.0, 75th percentile is 3.0

List: {1, 1, 3, 3, 5, 5}: 25th percentile is 1.5, 75th percentile is 4.5

List: {1, 2, 3, 3, 4, 4}: 25th percentile is 2.25, 75th percentile is 3.75

Next

New Round

Round 1:

You, along with everyone else in the room, are an applicant for a competitive program.

In this round, you will be assigned two random integers, A and B.

A is drawn from the set {1, 2, 3, 4, 5}.

B is drawn from the set {1, 3, 5}.

These two numbers, A and B, are the only two components of your endowment, and they will be displayed on your screen. **The sum of A and B (i.e., A+B) is your overall endowment. A is private information and is unknown to the program. B is public information and is known to the program.** You will choose whether or not to report A, having in mind that the program knows your B. If you choose to report A, you can only report the true number. Otherwise, you do not report, and the program will only have knowledge of B.

If you report, the program knows both A and B, and identifies your overall endowment perfectly. If you do not report, the program only knows B. Then, the program will form a belief about your A, and subsequently a belief about your overall endowment summing A and B. **The program's belief of A is positively correlated with its knowledge of B. In other words, the higher is your B, the higher is the chance that the program thinks your A is high.**

The program will admit half of the applicants with the highest overall endowment based on its knowledge or belief of each applicant's overall endowment. After all applicants make their reporting choices, the admission result will be displayed on your screen.

Next

Endowment and Reporting

Round 2:

Reminder 1: each applicant's endowment A is drawn from {1, 2, 3, 4, 5}, each applicant's endowment B is drawn from {1, 3, 5}. The program will admit half of all applicants based on its knowledge or belief about each applicant's overall endowment: A+B.

Admission statistics from last round:

Of all applicants that reported their endowments and were admitted:

the 25th percentile endowment A was 5.0, the 75th percentile endowment A was 5.0

Of all the applicants admitted, including those who reported their A and those who did not report their A:

the average value of endowment B was 5.0.

Reminder 2: 25th percentile is the value at which 25% of the endowments lie below that value; 75th percentile is the value at which 75% of the endowments lie below that value.

Your endowment for this round: Your Endowment A is 5, your Endowment B is 5.

Would you like to report your endowment A?

- report
- do not report

Next

Risk Aversion Elicitation

Instructions: In the following, you will face 10 decisions listed on your screen. Each decision is a paired choice between "Option A" and "Option B". You must choose between lottery A and lottery B, with lottery A delivering \$2 with certainty, and lottery B delivers \$1 with probability p and \$3 with probability (1-p). At the end of the survey, one random decision question will be picked and you will be paid the \$ amount that is provided by the lottery which you picked in that decision question (if you choose option A, you will be paid \$2 for sure; if you choose option B, you will be paid either \$1 or \$3 given their corresponding probabilities).

1. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 1/10 opportunity of \$1, 9/10 opportunity of \$3
2. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 2/10 opportunity of \$1, 8/10 opportunity of \$3
3. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 3/10 opportunity of \$1, 7/10 opportunity of \$3
4. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 4/10 opportunity of \$1, 6/10 opportunity of \$3
5. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 5/10 opportunity of \$1, 5/10 opportunity of \$3
6. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 6/10 opportunity of \$1, 4/10 opportunity of \$3
7. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 7/10 opportunity of \$1, 3/10 opportunity of \$3
8. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 8/10 opportunity of \$1, 2/10 opportunity of \$3
9. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 9/10 opportunity of \$1, 1/10 opportunity of \$3
10. Do you prefer option A or B?
   Option A: \$2 with certainty
   Option B: 10/10 opportunity of \$1, 0/10 opportunity of \$3

Bayesian-Update Simulation and Estimation Procedure

Here are the detailed simulation and estimation procedures for the Bayesian-updating process:

1. Define all possible school policies. We let $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$, $\gamma_0 \in \{0, 1, 2, 3, 4\}$, and $(\gamma_1, \gamma_2) \in \{(0,0), (0, 0.5), (0, 1), (0.5, 0), (1, 0), (0.5, 0.5)\}$. This gives us $5 \cdot 5 \cdot 6 = 150$ possible combinations of α, γ_0, γ_1, γ_2, which represent 150 unique school admission policies.

2. Simulate admission outcomes for each school policy. For each possible school policy (k), we run simulations with an applicant pool of 16 and 100 admission simulations for each possible $\{x, y\}$ pair appeared in the actual lab sessions. In total, there are $150 \cdot 16 \cdot 100 \cdot 29 = 6,960,000$ simulated individual admission outcomes. In each simulation,

   - each of the 16 applicants receive draws of $A = (1, 2, 3, 4, 5)$ and $B = (1, 3, 5)$ with the same rule as in the lab,
   - each applicant makes the comparison of $V_{R=1}$ vs. $V_{R=0}$, and chooses report with probability $\frac{\exp(V_{R=1})}{\exp(V_{R=1})+\exp(V_{R=0})}$,
   - once everyone receives a simulated reporting decision, the school determines the admission outcome for each applicant.

   Thus, for each school policy (k), we can calculate

   $$\Pr(O_t | A_t, B_t, R_t, x_{t-1}, y_{t-1}, k) = \frac{N^k_{A_t, B_t, R_t, O_t, x_{t-1}, y_{t-1}}}{N^k_{A_t, B_t, R_t, x_{t-1}, y_{t-1}}}$$

   where $N^k_{A_t, B_t, R_t, x_{t-1}, y_{t-1}}$ denotes the number of simulated observations with the exact combinations $\{A_t, B_t, R_t, x_{t-1}, y_{t-1}\}$, and $N^k_{A_t, B_t, R_t, O_t, x_{t-1}, y_{t-1}}$ denotes the number of simulated observations with the exact combinations $\{A_t, B_t, R_t, x_{t-1}, y_{t-1}, O_t\}$. In words, this is the probability that the admission outcome would be O_t given the initial admission statistics x_{t-1}, y_{t-1}, an individual's endowments A_t, B_t, reporting decision R_t, and school policy k.

3. Define the prior probability for each school policy. Define $\phi^{(1)} = \{\phi^{(1)}_k\}$ for $k = \{1, 2, .., 150\}$. For example, we assume the prior probability for each value of α is

   $$\{\phi^{(1)}_{\alpha=0}, \phi^{(1)}_{\alpha=0.25}, \phi^{(1)}_{\alpha=0.5}, \phi^{(1)}_{\alpha=0.75}, \phi^{(1)}_{\alpha=1}\}$$

   and similarly for γ_0 and (γ_1, γ_2), and all these priors are independent of each other. Then, the prior probability of policy k where $k = (\alpha = 0.5, \gamma_0 = 0, \gamma_1 = 0.5, \gamma_2 = 0.5)$ will be given by

   $$\phi^{(1)}_k = \Pr(\alpha = 0.5, \gamma_0 = 0, \gamma_1 = 0.5, \gamma_2 = 0.5) = \phi^{(1)}_{\alpha=0.5} \cdot \phi^{(1)}_{\gamma_0=0} \cdot \phi^{(1)}_{\gamma_1=0.5,\gamma_2=0.5}$$

An individual's prior probability is given by the vector $\phi^{(1)}$, where

$$\phi^{(1)} = \{\phi_k^{(1)}\} = \{\phi_1^{(1)}, \phi_2^{(1)}, ..., \phi_{149}^{(1)}, \phi_{150}^{(1)}\}$$

4. Perform Bayesian updating. For any subject-round observed in our real data, we observe

$$\{A_t, B_t, R_t, x_{t-1}, y_{t-1}, O_t\}$$

for $t = 1, ..., 50$. Then, we can calculate the posterior belief $\phi^{(t+1)}$ given the prior $\phi^{(t)}$ and the conditional probability $\Pr(O_t|A_t, B_t, R_t, x_{t-1}, y_{t-1}, k)$, which we calculated previously. The Bayesian updating equation for each policy k is:

$$\begin{aligned}\phi_k^{(t+1)} &= \Pr(k|A_t, B_t, R_t, x_{t-1}, y_{t-1}, O_t, \phi^{(t)}) \\ &= \frac{\Pr(O_t|A_t, B_t, R_t, x_{t-1}, y_{t-1}, k) \cdot \phi_k^{(t)}}{\sum_{k'=1}^{k'=150} \Pr(O_t|A_t, B_t, R_t, x_{t-1}, y_{t-1}, k') \cdot \phi_{k'}^{(t)}}\end{aligned}$$

5. Maximum likelihood. For a given set of parameters $(\phi^{(1)}, \Gamma)$, where Γ is a vector of coefficients for subject covariates, we can calculate the total log likelihood and find the maximum likelihood.

During the estimation process, we first perform a principal component analysis before adding covariates to the Bayesian update simulation. We keep the first nine components, which have eigenvalues larger than 1.

School Belief Derivation with an Extra Dimension

Suppose each student has real quality M in academic preparedness,

- with p_b probability $B_1 = M$, with $(1 - p_b)$ probability B_1 is randomly chosen from $\{1, 2, 3, 4, 5\}$ with equal probability
- with p_a probability $A = M$, with $(1 - p_a)$ probability A is randomly chosen from $\{1, 2, 3, 4, 5\}$ with equal probability

The school admits based on $B_2 + E(M|B_1, A)$. Now, we need to derive $E(M|B_1)$ and $E(M|B_1, A)$.

Test-Blind

Let $B_1 = x$. We know

$$P(B_1 = x|M = x) = p_b + \frac{1}{5}(1 - p_b)$$
$$P(B_1 = x|M = m \neq x) = \frac{1}{5}(1 - p_b)$$

then,

$$P(B_1 = x) = \frac{1}{5}\Big(p_b + \frac{1}{5}(1 - p_b)\Big) + \frac{4}{5}\Big(\frac{1}{5}(1 - p_b)\Big) = \frac{1}{5}$$

Then, we can write

$$P(M = x|B_1 = x) = \frac{P(B_1 = x|M = x) \cdot P(M = x)}{P(B_1 = x)} = p_b + \frac{1}{5}(1 - p_b)$$
$$P(M = m|B_1 = x) = \frac{P(B_1 = x|M = m) \cdot P(M = m)}{P(B_1 = x)} = \frac{1}{5}(1 - p_b)$$

Finally,

$$\begin{aligned} E(M|B_1 = x) &= x \cdot P(M = x|B_1 = x) + (15 - x) \cdot P(M = m|B_1 = x) \\ &= x\Big[\frac{1}{5} + \frac{4}{5}p_b\Big] + (15 - x)\Big[\frac{1}{5} - \frac{1}{5}p_b\Big] \\ &= 3 - 3p_b + p_b x \end{aligned}$$

For example, when $p_b = 0.5$, $E(M|B_1) = \frac{3}{2} + \frac{1}{2}B_1$.

Test-Required

Case 1: $A = B_1$**.** Suppose $A = B_1 = x$.

- Step 1. Write $P(A = B_1 = x|M = x)$.

– $P(A = B_1 = x|M = x) = \Big(p_a + \frac{1}{5}(1-p_a)\Big)\Big(p_b + \frac{1}{5}(1-p_b)\Big) = (\frac{1}{5} + \frac{4}{5}p_a)(\frac{1}{5} + \frac{4}{5}p_b)$

– $P(A = B_1 = x|M = m \neq x) = \frac{1}{5}\frac{1}{5}(1-p_a)(1-p_b)$

then,

$$P(A = B_1 = x) = \frac{1}{5}(\frac{1}{5} + \frac{4}{5}p_a)(\frac{1}{5} + \frac{4}{5}p_b) + \frac{4}{5}\frac{1}{5}\frac{1}{5}(1-p_a)(1-p_b)$$

- Step 2. Write $P(M = q|A = B_1 = x)$.

$$P(M = x|A = B_1 = x) = \frac{P(A = B_1 = x|M = x) \cdot P(M = x)}{P(A = B_1 = x)}$$
$$P(M = m|A = B_1 = x) = \frac{P(A = B_1 = x|M = m) \cdot P(M = m)}{P(A = B_1 = x)}$$

- Step 3. Write $E(M|B_1, A)$.

$$\begin{aligned} E(M|B_1, A) &= x \cdot P(M = x|A = B_1 = x) + \sum_{m \neq x} m \cdot P(M = m \neq x|A = B_1 = x) \\ &= x \cdot \frac{\frac{1}{5}(\frac{1}{5} + \frac{4}{5}p_a)(\frac{1}{5} + \frac{4}{5}p_b)}{\frac{1}{5}(\frac{1}{5} + \frac{4}{5}p_a)(\frac{1}{5} + \frac{4}{5}p_b) + \frac{4}{5}\frac{1}{5}\frac{1}{5}(1-p_a)(1-p_b)} \\ &\quad + (15 - x) \cdot \frac{\frac{1}{5}\frac{1}{5}\frac{1}{5}(1-p_a)(1-p_b)}{\frac{1}{5}(\frac{1}{5} + \frac{4}{5}p_a)(\frac{1}{5} + \frac{4}{5}p_b) + \frac{4}{5}\frac{1}{5}\frac{1}{5}(1-p_a)(1-p_b)} \\ &= \frac{x \cdot \frac{1}{5}(1 + 4p_a)(1 + 4p_b) + (15 - x) \cdot \frac{1}{5}(1-p_a)(1-p_b)}{1 + 4p_a p_b} \end{aligned}$$

Case 2: $A \neq B_1$**.** Take $A = 2$, $B_1 = 3$ as an example.

- Step 1. Write $P(A = 2, B_1 = 3|M = m)$ and $P(A = 2, B_1 = 3)$.
  – if $M = 2$, then

$$P(A = 2, B_1 = 3|M = 2) = \Big(p_a + \frac{1}{5}(1-p_a)\Big)\frac{1}{5}(1-p_b) = \frac{1}{5}(\frac{1}{5} + \frac{4}{5}p_a)(1-p_b)$$

  – if $M = 3$, then

$$P(A = 2, B_1 = 3|M = 3) = \frac{1}{5}(1-p_a)\Big(p_b + \frac{1}{5}(1-p_b)\Big) = \frac{1}{5}(\frac{1}{5} + \frac{4}{5}p_b)(1-p_a)$$

  – if $M = m \neq 2, 3$, then

$$P(A = 2, B_1 = 3|M = m) = \frac{1}{5}\frac{1}{5}(1-p_a)(1-p_b)$$

then,

$$P(A = 2, B_1 = 3) = \frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_a)(1-p_b) + \frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_b)(1-p_a) + \frac{3}{125}(1-p_a)(1-p_b)$$

- Step 2. Write $P(M = m|A = 2, B_1 = 3)$.

$$P(M = 2|A = 2, B_1 = 3) = \frac{P(A = 2, B_1 = 3|M = 2) \cdot P(M = 2)}{P(A = 2, B_1 = 3)}$$
$$P(M = 3|A = 2, B_1 = 3) = \frac{P(A = 2, B_1 = 3|M = 3) \cdot P(M = 3)}{P(A = 2, B_1 = 3)}$$
$$P(M = m|A = 2, B_1 = 3) = \frac{P(A = 2, B_1 = 3|M = m) \cdot P(M = m)}{P(A = 2, B_1 = 3)}$$

- Step 3. Write $E(M|B_1, A)$.

$$\begin{aligned} &E(M|B_1, A) \\ &\quad = \sum m \cdot P(A = 2, B_1 = 3|M = m) \\ &\quad = 2 \cdot \frac{P(A = 2, B_1 = 3|M = 2) \cdot P(M = 2)}{P(A = 2, B_1 = 3)} \\ &\qquad\quad + 3 \cdot \frac{P(A = 2, B_1 = 3|M = 3) \cdot P(M = 3)}{P(A = 2, B_1 = 3)} \\ &\qquad\quad + (1 + 4 + 5) \cdot \frac{P(A = 2, B_1 = 3|M = m) \cdot P(M = m)}{P(A = 2, B_1 = 3)} \end{aligned}$$

$$= \frac{2 \cdot \frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_a)(1 - p_b) + 3 \cdot \frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_b)(1 - p_a) + (1 + 4 + 5) \cdot \frac{1}{125}(1 - p_a)(1 - p_b)}{\frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_a)(1 - p_b) + \frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_b)(1 - p_a) + \frac{3}{125}(1 - p_a)(1 - p_b)}$$

when $A = a$, $B_1 = b$

$$= \frac{\frac{a}{25}(\frac{1}{5} + \frac{4}{5}p_a)(1 - p_b) + \frac{b}{25}(\frac{1}{5} + \frac{4}{5}p_b)(1 - p_a) + \frac{15-a-b}{125}(1 - p_a)(1 - p_b)}{\frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_a)(1 - p_b) + \frac{1}{25}(\frac{1}{5} + \frac{4}{5}p_b)(1 - p_a) + \frac{3}{125}(1 - p_a)(1 - p_b)}$$
$$= \frac{3(1 - p_a)(1 - p_b) + p_a a + p_b b - p_a p_b (a + b)}{1 - p_a p_b}$$