Voluntary Report of Standardized Test Scores: An Experimental Study *

Marty Haoyuan Chen[†]
University of Maryland

Ginger Zhe Jin[‡]
University of Maryland & NBER

July 2024

Preliminary, Comments Welcome

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, while subjects exhibit the ability to learn about the hidden school interpretations. Student subjects are more likely to hide their private information when the school is more lenient for non-reporting and when they receive a better draw on their other attribute observable to the school.

Using a structural model of student reporting behavior, we simulate admission outcomes for 16 counterfactual test-optional policies. Then, we examine the potential tradeoff between academic preparedness and diversity in a school's admission cohort. We find that test-blind is the worst in both dimensions, while test-required admits students with better academic preparedness and more diverse non-test attributes than 11 of the 16 test-optional policies. The few test-optional policies that are not dominated by test-required have more diversity but lower academic preparedness on average than test-required. This tradeoff occurs because our subjects do not possess perfect information of the school policy. When we simulate under perfect information, these two test-optional policies are also dominated by test-required in both dimensions.

Keywords: voluntary disclosure, unraveling, standardized test score, school admission, test-required, test-blind, test-optional.

JEL Codes: D8, D61, D63, I23, I24.

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. 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).[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]

---

[1]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.

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 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 $(A + B)$.

Our experiments, as well as a model of the simple admission game, highlight 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 strate-

4

gic reporting.

In addition to these two strategic incentives, we also 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. When we simulate under perfect information, these two test-optional policies are also dominated by test-required in both dimensions.

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.

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. Our model does not capture the cost of test preparation, and thus does not speak to the test-taking margin.

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

---

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

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 understand that private and public endowments are positively correlated in the applicant population, and every one'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:

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

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. Con-

ditional 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 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 subfigures, the square represents a uniform distribution of e and B in the applicant population. Subfigures (a), (b) and (c) present the admission outcome under test-required, test-blind, and test-optional, respectively. Subfigure (d) highlights the difference between test-required and test-blind; subfigure (e) highlights the difference between test-required and test-optional; and subfigure (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

---

[4]Note that this is different from the counter-signalling effect shown in Feltovich, Harbaugh and To (2002); Bederson et al. (2018). Counter-signalling is more complicated than reverse unraveling, because to ensure counter-signalling 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.
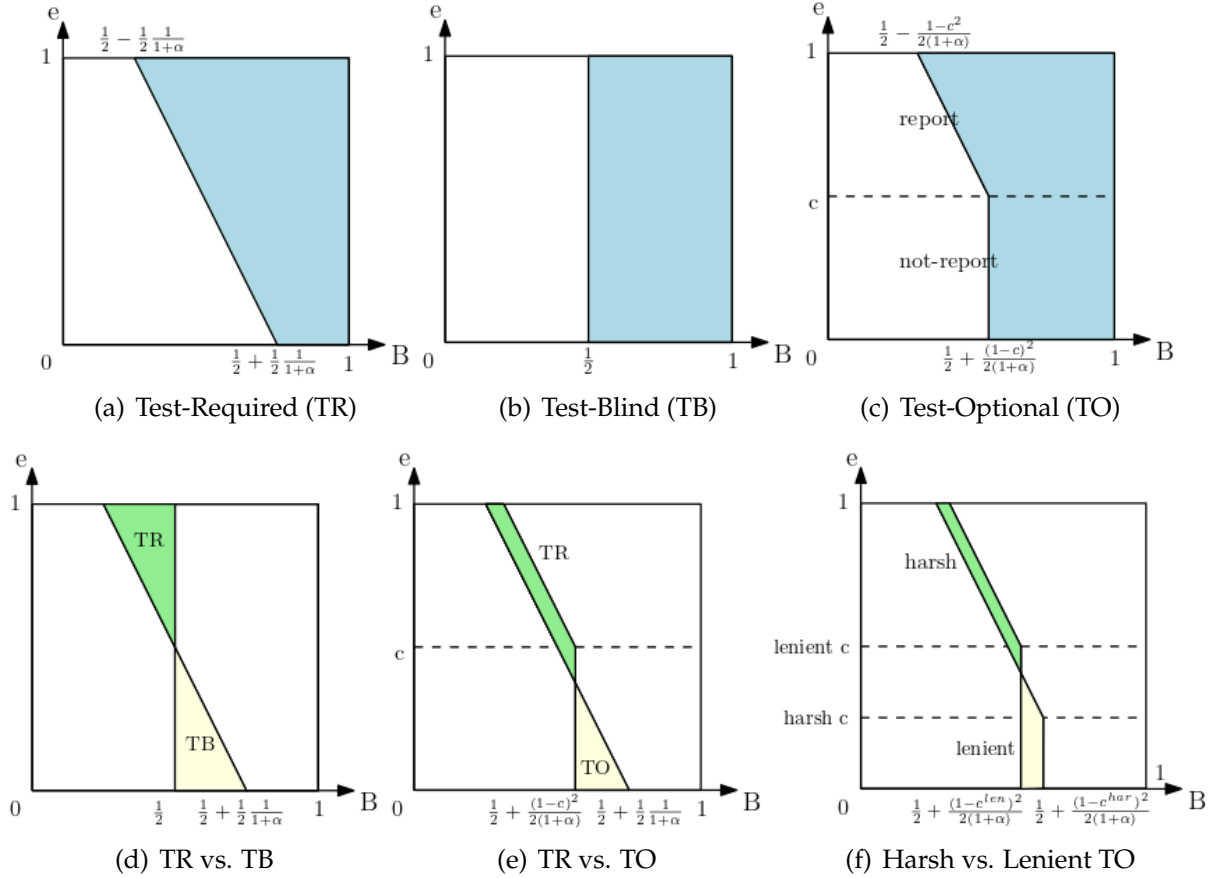
Figure 1: Composition of the Admission Cohort

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 subfigures (a), (b), and (c) represent admitted applicants. In subfigures (c), (d), and (e), we show the difference in admission composition between test policy pairs.

endowment and admit every student with a total endowment above it. This corresponds to the blue shaded area in Figure 1(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 Figure 1(b).

Putting test-required and test-blind in one graph, Figure 1(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 Figure 1(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-infererence 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 Figure 1(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.

Figure 1(e) 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.

Figure 1(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:

Test-required $\gg$ Harsh test-optional $\gg$ Lenient test-optional $\gg$ 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

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

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

$$\mathrm{E}_s[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] = \hat{\alpha} B_t + (1 - \hat{\alpha})[\hat{\gamma_0} + \hat{\gamma_1} x_{t-1} + \hat{\gamma_2} y_{t-1}]$$

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, $\max_{R_t \in \{0,1\}} \mathrm{EU} = p_t U^a + (1 - p_t) U^r$. 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

$$\underbrace{f(B_t + A_t)}_{\text{Admission prob. from reporting}} > \underbrace{\hat{\alpha} f(B_t + B_t) + (1 - \hat{\alpha}) f(B_t + \hat{\gamma_0} + \hat{\gamma_1} x_{t-1} + \hat{\gamma_2} y_{t-1})}_{\text{Expected admission prob. from non-reporting}}$$

$$= f(B_t + \mathrm{E}_s[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] - \pi(\rho))$$

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:

$$R_t = \mathbb{1}\left\{ A_t > \mathrm{E}_s[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] - \pi(\rho) \right\} \tag{6}$$

14

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). 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 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. [5]

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

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

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

---

[5]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.

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 modelling 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.

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.

$$R_t = \mathbb{1}\Big\{ A_t > \mathrm{E}[\bar{A}_t | B_t, x_{t-1}, y_{t-1}, R_t = 0] \Big\} \tag{7}$$

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 have risk aversion but the computer does 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 the risk measure of each subject, (iii) by introducing a probability of irrationality, and (iv) by only focusing on later-round results.

	T1	T2	T1-T2	p-value
	(1)	(2)	(3)	(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.

4.1 Trends in 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. Subfigures (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 simu-

(a) T1: lab

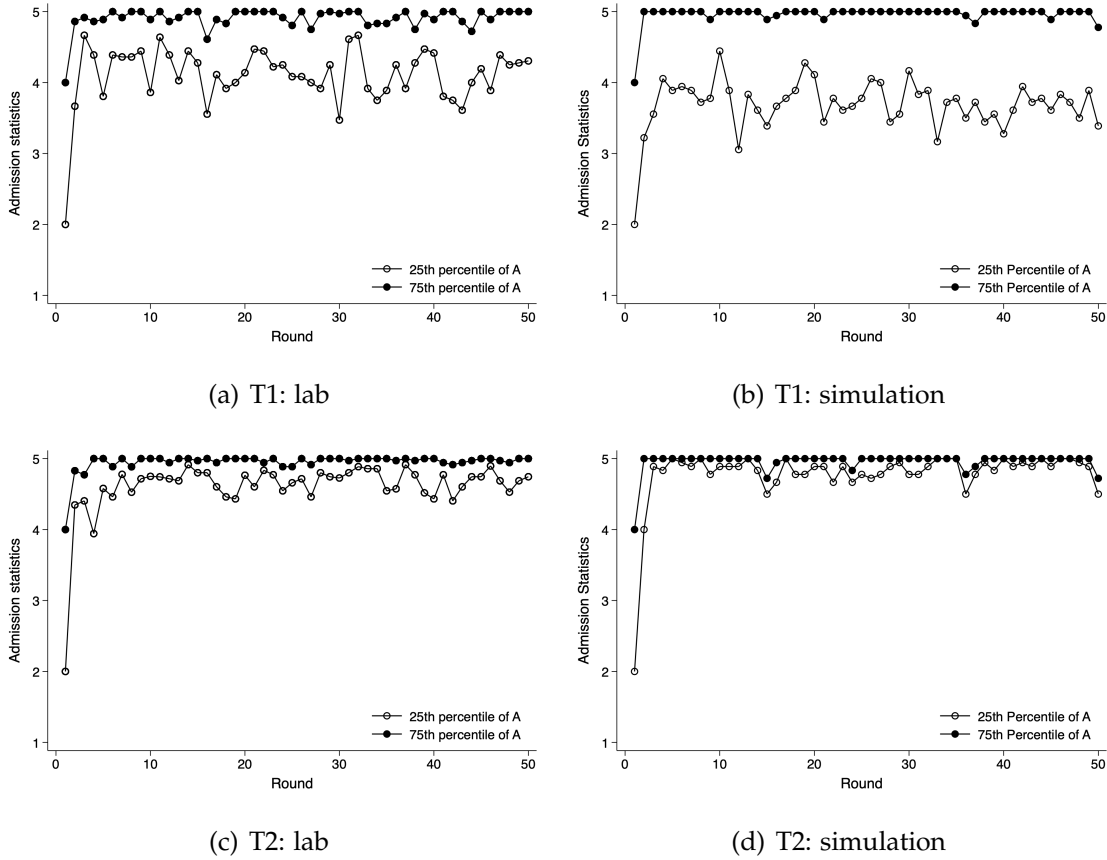(b) T1: simulation

(c) T2: lab

(d) T2: simulation

Figure 2: Trends in Admission Statistics

Note: This figure shows the trends in admission statistics by round. Subfigures (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. Subfigures (b) and (c) present the trends when we simulate subject reporting decision with perfect information under T1 and T2, respectively.

lated 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 "converging" 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.

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.

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
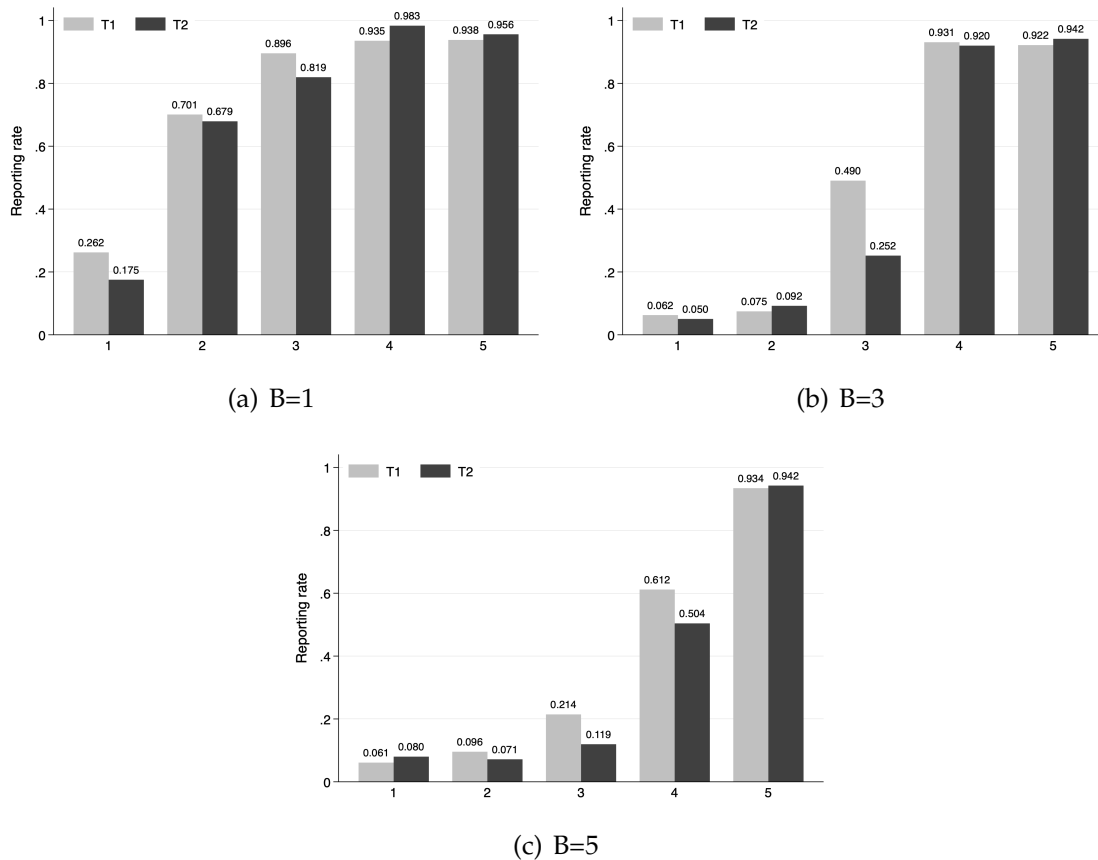
Figure 3: Reporting Rates in Rounds 21-50

Note: This figure presents the reporting rates by public and private endowments in the last 30 rounds of the experiment. Each subfigure represents one possible public endowment (B) and the horizontal axis in each subfigure 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.

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. [6]

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 subjecst 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 was 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 7. 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.

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 7 and the actual decision rule given by Equation 6, 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

---

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

Table 3: Hypothetical Reporting Rule with Perfect Information

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

[*] *Notes:* This table reports the hypothetical reporting rule (indicated by Equation 7) 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.
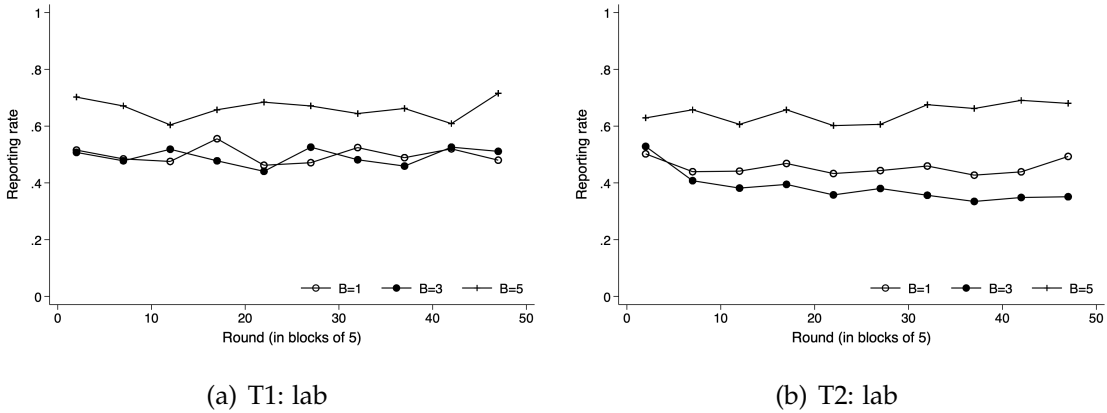

(a) T1: lab   (b) T2: lab

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.

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. In general, Figure 4 shows no clear evidence of 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 Section 5.
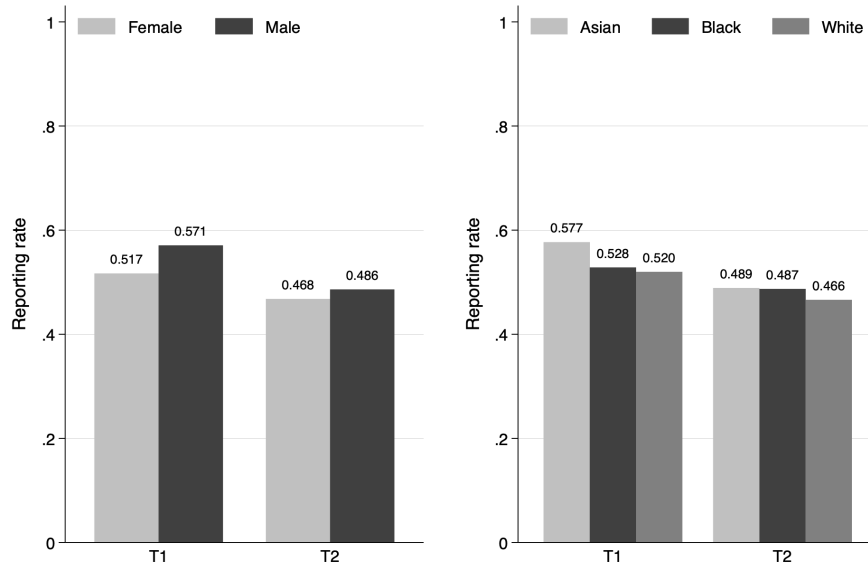
Figure 5: Reporting Rates by Gender and Race

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.

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

$$V_{R_t=1} = A_t$$
$$V_{R_t=0} = E_s(\bar{A}_t) - \pi(\rho) + \epsilon$$
$$= \hat{\alpha} \cdot B_t + (1 - \hat{\alpha})[\hat{\gamma}_0 + \hat{\gamma}_1 \cdot x_{t-1} + \hat{\gamma}_2 \cdot y_{t-1}] - \delta \cdot \rho + \epsilon$$

where ϵ is the logit error. As discussed previously, we used the multiple price list method to elicit risk preference of each subject. ρ represents the question number at which each subject switched from getting a lottery to getting a sure prize. A higher ρ suggests higher risk aversion. Given the utilities, a rational subject would report in period t if $V_{R_t=1} > V_{R_t=0}$.

To capture subject learning at any point in the experiment, 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. Since subjects in different treatments were told the same instructions, learning was also the major distinction of each treatment specification.

The parameters of the discrete model were estimated using maximum likelihood. The results are provided in Table 4. We set the coefficient of A_t in the logit model to always be one, 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. [7] 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. The coefficient for the risk aversion index also became statistically indistinguishable from

---

[7]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.

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)
Risk Aversion Index	-0.03**	-0.04**	-0.04	
	(0.01)	(0.02)	(0.03)	
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. The learning variables in Columns 3 and 4 include 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 all previous rounds, subject acceptance rate pooled and by reporting decision under the same public endowment all previous rounds. The subject covariates in Column 4 include those summarized in Table 1. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

zero. A significant negative coefficient would suggest that more risk averse subjects are less likely to report, contradicting our hypothesis. However, it is still surprising that the coefficients remain negative across all models. One explanation is that, while the model assumes reporting implies certainty and non-reporting implies uncertainty, in reality a subject would still face the uncertainty of not knowing the school's admission policy and other applicants' reporting decisions even if she reported. 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 0.72-0.75. Thus, subjects tend to overweight their public endowment and underweight school attributes when they make reporting decisions.

The predicted subject reporting rates from these model specifications 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 and 1.4 for model 4), 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 predictions are 35.1% or 33.9%; the actual reporting rate in T2 when the private endowment was 4 was 80.5%, while the model predictions are 76.6% or 76.2%. In general, the model fit is better in T1 than in T2. [8]

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:

$$\text{Rule of thumb 1:} \quad R_t|\text{naive} = 1 \text{ if } A_t \in \{4, 5\}$$
$$\text{Rule of thumb 2:} \quad R_t|\text{naive} = 0 \text{ if } A_t \in \{1, 2, 3\}$$

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

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

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 models 3 and 4, we estimate θ_1 to be 3% and θ_2 to be 7-8%, which are small but non-trivial. The average distances for models 3 and 4 are reduced by roughly one third 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

---

[8]To better capture subject learning and better predict subject reporting decision, we are exploring a Bayesian learning model in which subjects form a belief about the probability of each possible school interpretation formula, and update their belief at the end of each round based on the observed admission statistics and their admission outcome. This would also allow us to relax the assumptions we make when we later perform counterfactual simulations.

Table 5: Structural Model Predicted Reporting Rates

Panel A:
Reporting rate (%)

	All				Treatment: T1			Treatment: T2		
	Actual	Model (2)	Model (3)	Model (4)	Actual	Model (3)	Model (4)	Actual	Model (3)	Model (4)
A=1	17.9	19.1	17.5	18.0	21.0	19.7	20.7	14.6	15.2	15.2
A=2	28.3	29.8	25.1	25.3	28.8	26.9	26.9	27.7	23.1	23.6
A=3	40.6	51.8	42.8	42.6	50.6	50.4	51.1	30.4	35.1	33.9
A=4	81.6	72.7	80.2	80.1	82.7	83.7	83.8	80.5	76.6	76.2
A=5	93.8	82.8	93.5	93.2	93.3	93.8	93.4	94.4	93.2	93.1
Avg. distance (unweighted)		6.8	1.5	1.4		1.0	0.8		3.0	2.8
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)	Model (4)	Actual	Model (3)	Model (4)
A=1	17.9	16.4	17.0	17.5	21.0	19.1	20.1	14.6	14.8	14.7
A=2	28.3	25.6	24.3	24.6	28.8	26.1	26.1	27.7	22.4	22.9
A=3	40.6	44.6	41.6	41.3	50.6	48.9	49.5	30.4	34.1	32.9
A=4	81.6	86.6	81.6	81.7	82.7	84.8	85.1	80.5	78.2	78.1
A=5	93.8	91.6	93.9	93.8	93.3	94.2	93.9	94.4	93.7	93.6
Avg. distance (unweighted)		3.1	1.2	1.0		1.9	1.5		2.4	2.1
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. Both pooled reporting rates and reporting rates by treatments are presented. 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.

subject fixed effects than when we add subject attributes. This is because subject fixed effects capture more subject characteristics that predict reporting decisions than what we observe from survey responses.

6   Welfare Implications

6.1   Academic Preparedness and Diversity Tradeoff

In this section we discuss the tradeoff between academic preparedness and diversity of a school's entering cohort. Assuming that schools always admit the top candidates (i.e. candidates with the highest overall endowment) until they reach fully 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 less 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 functions 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 parame-

29

Table 6: List of School Interpretations of Non-Reporting

Setting	α	γ_0	γ_1	γ_2	$E[A_t|R_t = 0]$
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.

ters 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.

The full list of simulated counterfactual school interpretations is presented in Table 6. To present the academic preparedness and diversity tradeoff, 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. In other words, the college objective functions 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
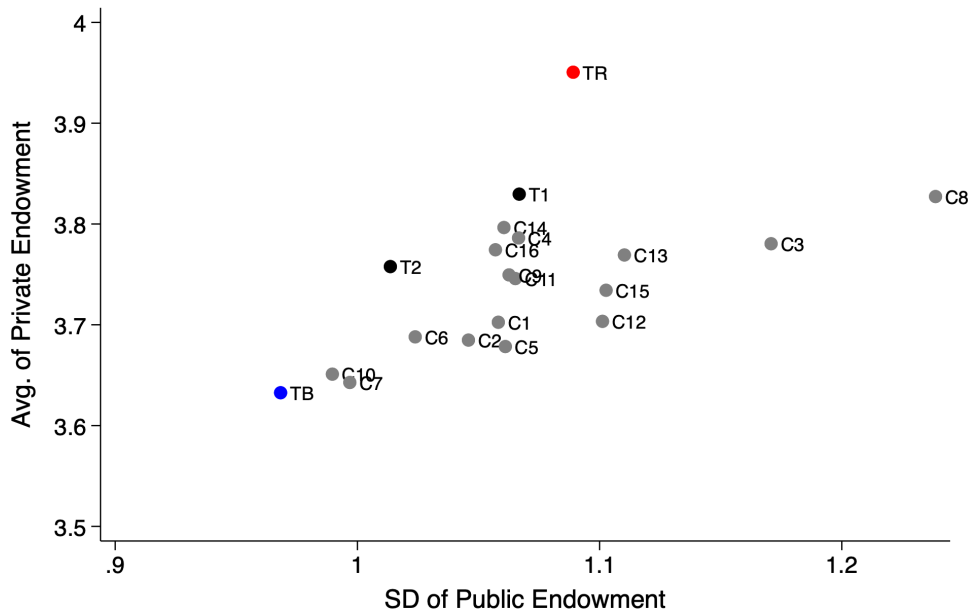
Figure 6: Academic Preparedness and Diversity Tradeoff

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).

status. [9]

Figure 6 shows the tradeoff illustration 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. First, there does not appear to be much of a tradeoff: school policies that admit students with higher quality also admit students from a more diverse background. 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. Voluntary reporting raises a lesser degree of the same concern. Compared to a test-required policy, it makes standardized test performance less visible, and forces

_____

[9]In this section, we use "private endowment" interchangeably with "standardized test scores" or "SAT/ACT scores", and use "public endowment" interchangeably with "non-test attribute".

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.

A 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 cohort in non-test attributes. While test-blind and test-optional schools are limited in their ability to do so, a test-required school provides students with a weak non-test attributes an opportunity to stand out through standardized tests. On the contrary, a test-blind policy is dominated by all other policies plotted in Figure 6 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.

Eleven 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 7 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 6), we assume the same naivety probabilities as those reported in Table 5, Column 3. Comparing Figures 6 and 7, 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 6 and 7? 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
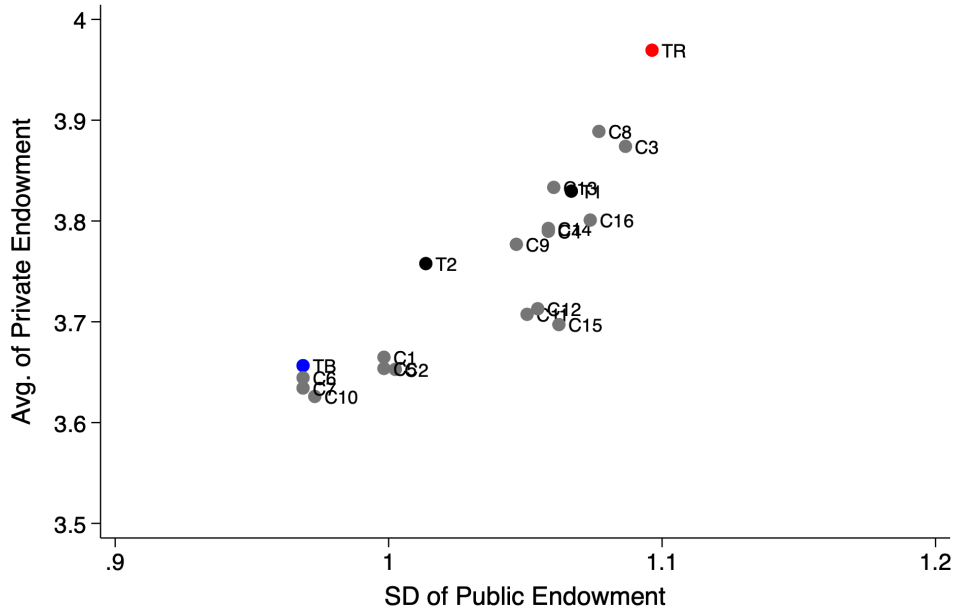
Figure 7: The Tradeoff with Perfect Information

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.

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 6. 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-
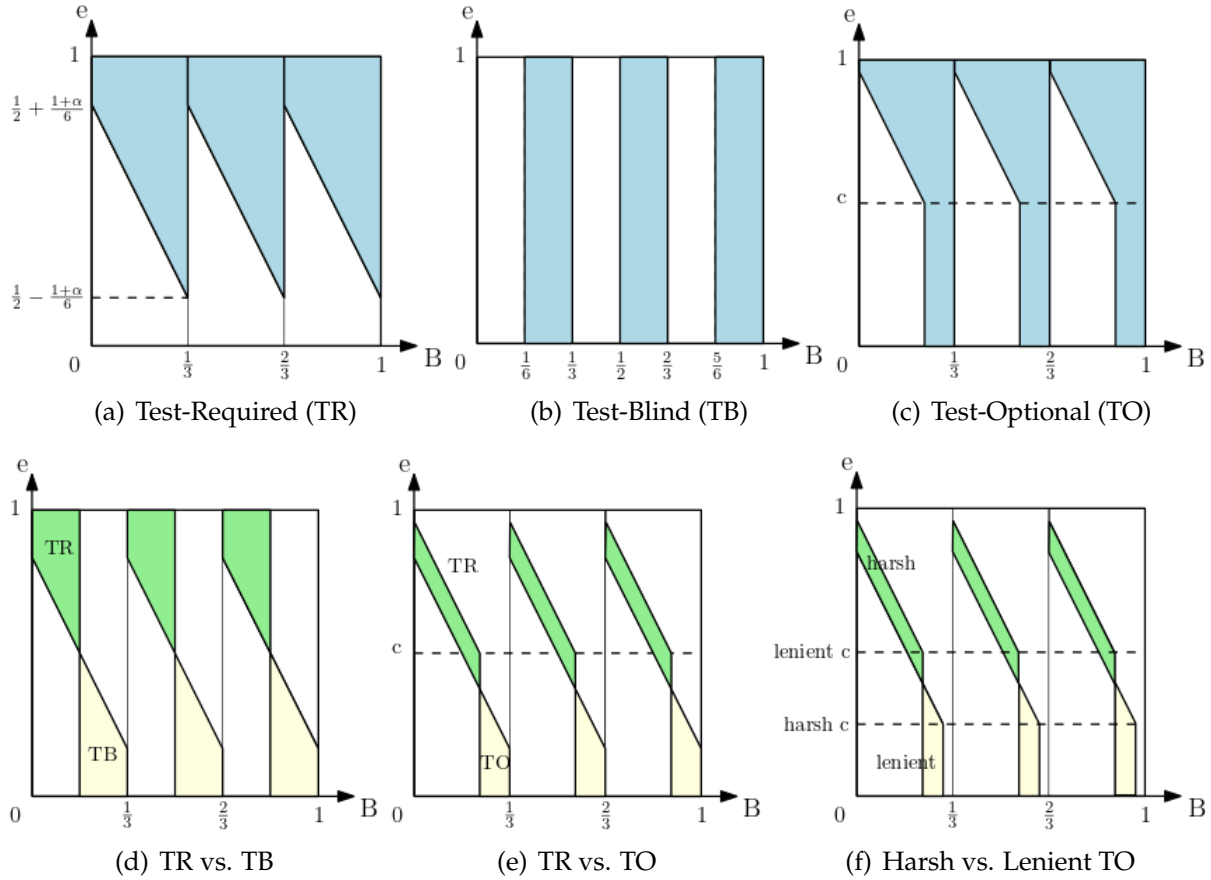
Figure 8: Composition of the Admission Cohort

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 subfigures (a), (b), and (c) represent admitted applicants. In subfigures (c), (d), and (e), we show the difference in admission composition between test policy pairs.

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 Imposing Admission Quota by Public Endowment

The results shown in Figure 6 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 8 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. The analysis can be extended to school policies with alternative forms of quotas, but they send similar messages.

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. 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.

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. The same results can be extended to schools that impose admission quotas by non-test attributes.

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. Voluntary reporting appears to be a wrong tool for diversity.

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.

Chetty, Raj, David J. Deming, and John N. Friedman. 2023. "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges." NBER Working Paper #31492.

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.

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.

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.

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.
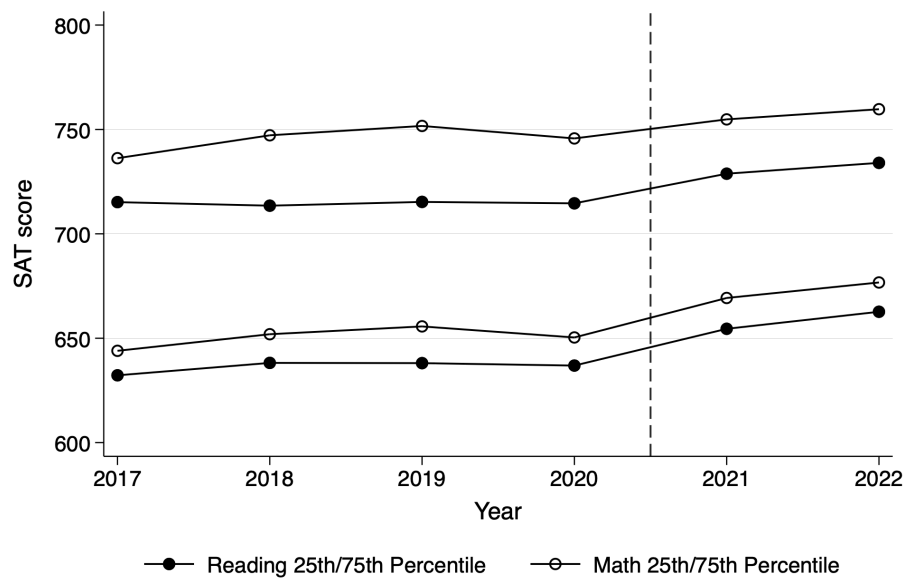
A   Appendix: Supplementary Materials

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.
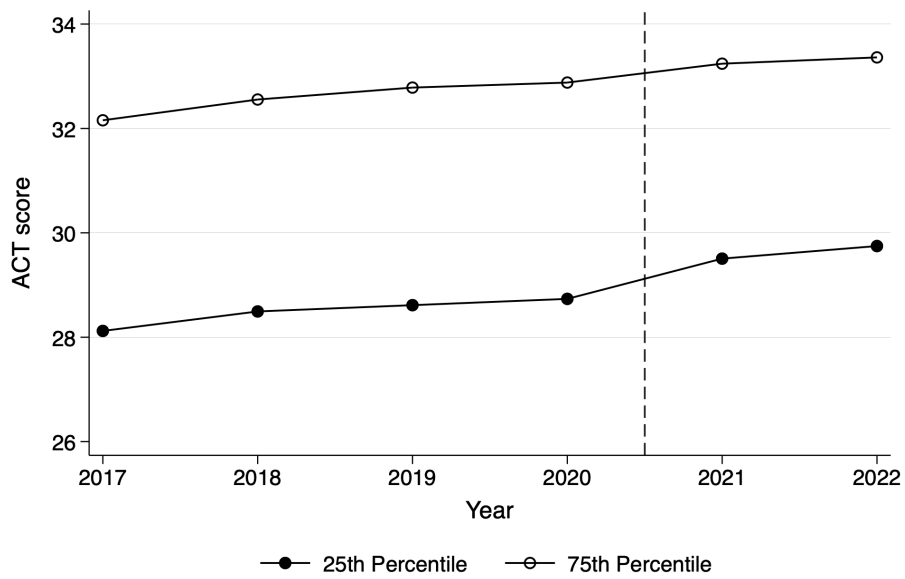
Figure A.2: Trends in Reported ACT Scores

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.
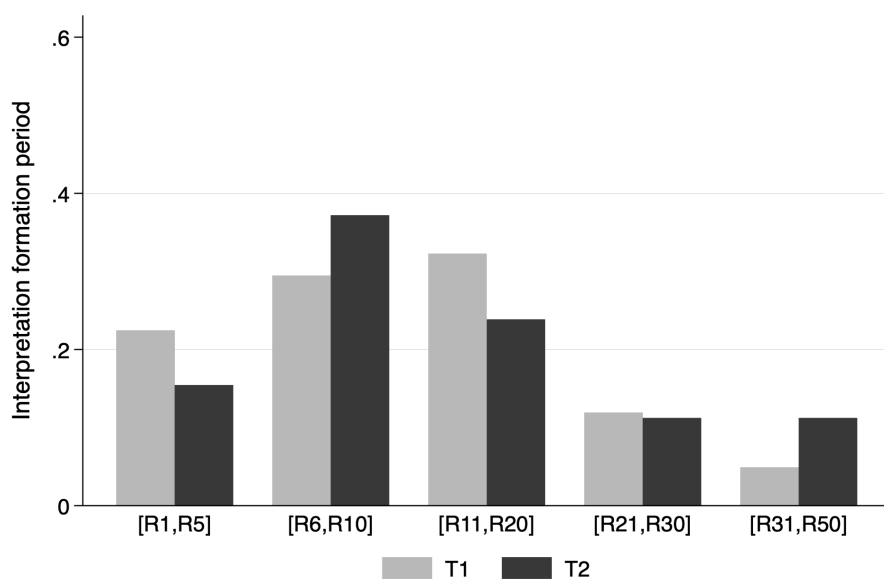
Figure A.3: Time of Belief Formation

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?".

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.

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