Dropping Standardized Testing for Admissions Trades Off Information and Access

Nikhil Garg
Cornell Tech and Technion
ngarg@cornell.edu

Hannah Li
Stanford University
hannahli@stanford.edu

Faidra Monachou
Stanford University
monachou@stanford.edu

September 7, 2021

Abstract

We study the role of information and access in capacity-constrained selection problems with fairness concerns. We develop a theoretical framework with testable implications that formalizes the trade-off between the (potentially positive) informational role of a feature and its (negative) exclusionary nature when members of different social groups have unequal access to this feature. Our framework finds a natural application to recent policy debates on dropping standardized testing in college admissions.

Our primary takeaway is that the decision to drop a feature (such as test scores) cannot be made without the joint context of the information provided by other features and how the requirement affects the applicant pool composition. Dropping a feature may exacerbate disparities by decreasing the amount of information available for each applicant, especially those from non-traditional backgrounds. However, in the presence of access barriers to a feature, the interaction between the informational environment and the effect of access barriers on the applicant pool size becomes highly complex. In this case, we provide a threshold characterization regarding when removing a feature improves both academic merit and diversity. Finally, using application and transcript data from the University of Texas at Austin, we illustrate that there exist practical settings where dropping standardized testing improves or worsens all metrics.

1 Introduction

Recent debates on the use of standardized testing in college admissions have increasingly garnered national attention, particularly during the COVID-19 pandemic as test centers shut down and schools were forced to reconsider their admissions practices (Anderson, 2020). Independently of the COVID-19 pandemic, in an attempt to increase equity and diversity in admissions, the University of California (UC) recently settled a lawsuit by eliminating all consideration of SAT and ACT scores for admissions and scholarships through 2025, following an earlier decision to suspend testing requirements and ultimately design its own test (Nieto del Rio, 2021).

These discussions primarily center on highly selective institutions and their efforts to shape the student body through the admissions process.[1] These schools promise great opportunities to their students, but – due to perceived capacity constraints – limit their acceptances to students that they deem to have high potential in academics, athletics, creative endeavors, or leadership and service (Espenshade and Radford, 2013). They typically attempt to identify these students through a combination of standardized tests, high school grades, letters of recommendation, personal essays, and extracurricular activities (Espenshade and Radford, 2013; Zwick, 2002).

The question is whether each of these components, and the application as a whole, allows the schools to assess individuals from different backgrounds effectively and 'fairly,' including students from different racial, ethnic, and socioeconomic groups. Implicitly, the debate concerns how to *design* an admission policy to aid fair and efficient decision-making, in terms of both deciding which information to collect from applicants and how to use the information. Our exposition focuses on the context of college admissions; however, our model and the questions we ask are more broadly applicable to other settings of information design and fair decision-making in capacity-constrained settings, such as labor markets, award committees, and social welfare programs.[2] In each of these cases, the decisions are being made on limited information but have far-reaching consequences for employment or education opportunities. Thus, it is important to analyze these policies and their potential disparate impact across different groups of applicants.

Background. The high-profile debates surrounding the use of standardized testing for admissions, social scientists and education experts have highlighted specific fairness concerns. Test critics argue that tests exhibit racial gaps (Reardon, 2011) and reinforce inequality in higher education (Reeves and Halikias, 2017). Espenshade and Radford (2013) find that only 8% of lower-income compared to 78% of high-income students use a test preparation service. The testing process is expensive and time-consuming, discouraging low-income students from even applying. (Hyman, 2016) finds that "for every ten poor students who score college-ready on the ACT or SAT, there are an additional five poor students who would score college-ready but who take neither exam" and so cannot apply

---

[1]Most schools are not selective, accepting most applicants. The admissions considerations of these schools differ substantially from those of more selective institutions (Selingo, 2020).

[2]Our model and insights apply in settings where there is a trade-off in the value of additional information and the fraction of applicants who can provide it. For example, in means-testing welfare programs, requiring long forms might help in better targeting benefits but also might discourage eligible recipients from applying (Hernanzi et al., 2004).

to colleges that require it.[3]

Supporters of testing argue that it is "a systematic means of collecting information," and so helps decision-making when used appropriately (Phelps, 2005). Some supporters claim that tests actually *help* schools evaluate under-represented minorities; in the absence of standardized testing, "a capable student from a little-known school in the South Bronx may be more challenging to evaluate," further benefiting students from privileged – and historically familiar – backgrounds (Bellafante, 2020). A report released by University of California further states that test scores are more accurate for minority groups, and "that consideration of test scores allows greater precision when selecting from [under-represented minority] populations" (University of California Standardized Testing Task Force, 2020). Other application components such as recommendation letters (Dutt et al., 2016) and application essays (Alvero et al., 2021) may also be unreliable.[4] A school that does not consider test scores must rely more heavily on these components.

The competing claims from critics and supporters largely center around two issues: *access* and *information*. We capture these arguments in favor of and against dropping test scores and formalize the underlying trade-off. The model considers a Bayesian school that wishes to admit students based on their skill level, which we refer to as "academic merit," and also values the "diversity" of the admitted class. The school admits students to meet a capacity constraint and tries to maximize the average academic merit of the accepted cohort. However, it has imperfect knowledge of the students skills and instead must rely on noisy and potentially biased signals, one of which is the test score. The school can then choose whether or not to require the test score. We evaluate the school's admission policy in terms of the average academic merit, overall diversity, and individual fairness of the students admitted (i.e., how the policy affects students of different groups and skill levels).

We focus on the trade-off between two effects:

Differential informativeness. Colleges often have better information – through, e.g., familiar letter writers and transcripts – on students from privileged backgrounds, and so can better estimate their true academic merit. Standardized testing reduces this measurement gap, and so in particular helps colleges identify well-qualified, non-traditional students.

Disparate access and applicant pool composition. Some students – especially those from disadvantaged backgrounds – either do not take standardized tests or do not report their scores,[5] due to cost and other access barriers. Without a test score, students cannot apply

---

[3]After eliminating GRE requirements, UC Berkeley saw an 82% increase in the number of under-represented minority applicants to master's programs in the 2020-2021 cycle: "while overall graduate applications have increased 19 percent when compared to [the 2019-2020 cycle], the number of underrepresented minority (URM) doctoral applicants increased by 42 percent and URM applicants to academic master's programs increased by 82 percent" (Aycock, 2021).

[4]For example, letter writers use different language to describe women and other under-represented groups, giving weaker recommendations (Dutt et al., 2016), and application essays have a stronger correlation to reported household income than do SAT scores (Alvero et al., 2021) (although they are not necessarily differentially scored).

[5]A University of California report on testing states that under-represented students might be discouraged from applying based on their score, even if their score would be competitive (University of California Standardized Testing Task Force, 2020).

to a school with a test requirement, even if they are well-qualified. Dropping the requirement thus expands the applicant pool but also alters its composition at different rates across groups.

Contributions. Given these effects, we study: *Under what settings of informativeness and disparate access should standardized testing be dropped from admissions, if a college values both diversity and academic merit?* Our technical contributions are as follows.

1. We introduce a Bayesian model with *multiple* application components that allows us to study the *design* of the information structure used in the application process. We formalize a trade-off between informativeness and access, two basic arguments in favor of and against the inclusion of a given feature, and show how the set of features required influences the academic merit and diversity through these two competing effects. Our main technical insight shows that differences in the total *variance* of features lead to information disparities across groups: even though the school manages to correct for the existing mean bias in the features of different groups, it is generally impossible to correct for variance.

   Our model leverages the statistical discrimination framework developed in Phelps (1972) and Arrow (1971), which we extend by considering multiple features and access disparities. Within this framework, we define two fairness notions: *diversity* and *individual fairness*. The former captures disparities at a group level. The latter quantifies disparities in individual opportunities, by measuring the difference in the admissions probability between two individuals of equal skill but different demographic characteristics.

2. We provide a testable framework for evaluating the different trade-offs that arise in these decisions. Using application and transcript data from the University of Texas at Austin, we demonstrate how an admissions committee could measure the trade-off in practice to better decide whether to drop their test scores requirement. We show that there exist practical settings both in which dropping testing worsens or improves all metrics.

Our primary takeaway for practice is that the decision to drop testing cannot be made without jointly considering the interaction between the information provided by other features *relative to test scores* and the rate at which dropping the test requirement affects the applicant pool composition. This interaction between information and access is complex. In the absence of access barriers to the test, the information loss incurred by dropping the test requirement always decreases academic merit, but has an ambiguous impact on diversity. However, in juxtaposition with unequal access barriers, the informational disparities can be amplified or reduced by the effect that the expanded access has on the applicant pool composition. We characterize the settings where dropping test scores introduces a trade-off between diversity and academic merit and where it improves or worsens all objectives.

More broadly, our work further provides a Bayesian framework for predicting the effect of adding new features into the admissions process. Given some knowledge of the informativeness and access barriers associated with the new feature, the model can be used to reason about how the new

feature would interact with the current set of features. We thus believe that our work provides a useful conceptual framework for studying emerging problems in fair decision-making and public policy.

1.1 Related Work

Our work broadly relates to the study of discrimination and admissions in the economics and fair machine learning communities.

Economics of discrimination. Many economics works study affirmative action in admissions (Abdulkadiroğlu, 2005; Avery et al., 2006; Chade et al., 2014; Chan and Eyster, 2003; Epple et al., 2006; Fershtman and Pavan, 2020; Fu, 2006; Immorlica et al., 2019; Kamada and Kojima, 2019), and more broadly discrimination in markets (Coate and Loury, 1993; Fang and Moro, 2011; Foster and Vohra, 1992; Lang and Manove, 2011; Levin, 2009; Temnyalov, 2018). In contrast to taste-based discrimination theories (Becker, 1957), statistical discrimination theory (Arrow, 1971; Phelps, 1972) shows that group differences can arise in equilibrium even if groups are *ex ante* identically skilled. In particular, in Phelps' seminal Gaussian framework (Phelps, 1972), employers receive a noisy signal about each worker's skill and are incentivized to use information about a group's belonging to infer the true skill. This exact statistical discrimination approach is surprisingly rare in the admissions literature (except Emelianov et al. (2020); Kannan et al. (2019)). To the best of our knowledge, our work is the first to extend Phelps' model to multiple features with non-identical distributions and incorporate access asymmetries to some feature.

Fairness in machine learning and mechanism design. Recent machine learning work applies fairness notions to college admissions and related allocation problems (Cai et al., 2020; Emelianov et al., 2020; Faenza et al., 2020; Haghtalab et al., 2020; Hu et al., 2019; Immorlica et al., 2019; Kannan et al., 2019; Kleinberg and Mullainathan, 2019; Liu et al., 2020; Mouzannar et al., 2019).

In a single-feature setting, several works analyze admissions or hiring decisions when evaluation of one group is noisier than another (Emelianov et al., 2020; Fershtman and Pavan, 2020; Temnyalov, 2018). Most related, Emelianov et al. (2020) study how differential variance of a single feature affects the admissions decisions of a school that greedily admits students with the highest test scores, without factoring in the differential variance; they show that affirmative action can improve both diversity and academic merit. In contrast, our work studies the impact of differential bias and variance when students have *multiple* features and schools can potentially drop a feature.

Another line of literature considers different types of barriers, including implicit bias (Faenza et al., 2020), downstream effects of school admissions in later employment (Kannan et al., 2019), and when only one group can take the test multiple times (Kannan et al., 2021). These barriers affect the treatment of applicants, but do not prevent students from even applying, as is our focus.

Finally, a follow-up paper (Liu and Garg, 2021) builds upon our work to provide (im)possibility results under test-optional policies.

2　Model

We develop a model where the school can design their admissions procedure and, in particular, choose the information that it requires the applicants to submit.

We consider a continuum of students and a single school. A unit mass[6] of students is applying to college. Each student belongs to a group $g \in \{A, B\}$, and the mass of students in group B is π. Each student has a latent (unobserved) *skill level* q, normally distributed according to $\mathcal{N}(\mu, \sigma^2)$ identically for each group, as well as a set of observed *features* $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$. Each θ_k is a noisy function of q, i.e., $\theta_k = q + \epsilon_k$, $k = 1, \ldots, n$, with Gaussian noise $\epsilon_k \sim N(\mu_{gk}, \sigma_{gk}^2)$. The distribution of noise ϵ_k is feature- and group-dependent, but each ϵ_k is drawn independently across features and students. Features represent application components like recommendation letters, grades, and test scores.

Students differ in their *access* to the features: only a fraction γ_g of group $g \in \{A, B\}$ has access to the *full* set of features FULL $= \{1, \ldots, K\}$, i.e., $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$; the remainder only has access to the *subset* SUB $= \{1, \ldots, K-1\}$. Whether a student has access to all features is independent of skill q and conditionally independent of the feature values given group membership. We assume that when a student does not have access to feature K, then they cannot apply to a school that requires it.

Admissions policy

We now turn to the question of interest: the design of the admissions policy. The school admits a mass C of students. The school's admissions procedure consists of a feature requirement policy, skill estimation, and then selection given estimates.

The feature requirement policy choice is whether to require the full set of features or the subset. If the school requires the full set, then students without full access cannot apply. If it only requires the subset, then it observes only that subset for each student.

Then, given a student's features $\boldsymbol{\theta}$, the school estimates a *perceived skill* $\tilde{q}$ of their true skill q. The school is Bayesian, knows the distribution of q and the (group-dependent) distributions of ϵ_k, and is *group-aware*: it can use the student's group membership in constructing its estimate.[7] The resulting Bayesian estimate is the 'best' one can do, given the available information:

$$\tilde{q}(\boldsymbol{\theta}, g) \triangleq \mathbb{E}[q \mid \boldsymbol{\theta}, g].$$

After estimating the skill level of each applicant, the school selects the mass C of students with the highest skill estimates $\tilde{q}$. This selection process induces a *threshold* $\tilde{q}_S^*$ such that applicants with perceived skill above the threshold are admitted. (In Section 4.3 we also study selection policies

---

[6]For exposition clarity, we describe the characteristics of individual students. Such statements should be interpreted as illustrative of the corresponding continuum system.

[7]Ignoring group attributes is an oft-proposed but often problematic policy proposal to combat bias in machine learning tasks (Corbett-Davies and Goel, 2018). We evaluate *group-unaware* estimation in Appendix A.1.

utilizing *affirmative action*, where the school uses potentially group-dependent thresholds.[8])

Holding the estimation and selection policies fixed (except in Section 4.3), the *admissions policy* is determined by the feature requirement decision. Let P_S denote the admissions policy requiring feature set S.

Academic merit and fairness metrics

We evaluate a policy P using three metrics on the admitted class. Let $Y \in \{0, 1\}$ denote the admission decision for a given student; $Y = 1$ means that the student is admitted.

Academic merit $\mathbb{E}[q \mid Y = 1, P]$, the expected skill level of accepted students. We also use group-specific measures, $\mathbb{E}[q \mid Y = 1, g, P]$.

Diversity level $\tau(P)$, the fraction of students admitted that are of group B. Policy P satisfies *group fairness* if and only if the admission fraction matches the population, i.e., $\tau(P) = \pi$.

Individual fairness gap $I(q; P)$, the difference in admissions probability between two students of identical true skill q, one belonging to group A and the other to group B:

$$I(q; P) \triangleq \mathbb{P}\left(Y = 1 \mid q, A, P\right) - \mathbb{P}\left(Y = 1 \mid q, B, P\right).$$

Policy P satisfies *individual fairness* if and only if the gap is 0 for all skill levels q.

We characterize these three metrics as they depend on the policy P and the model parameters, as well as how they trade off with one another.

College admissions and relationship to practice

While our model and results are more general, our exposition primarily considers undergraduate college admissions in the United States and the debate to drop standardized testing as our main running example. We focus on how policies differentially affect privileged (group A) versus disadvantaged (group B) students.

We refer to the potentially inaccessible last feature, $k = K$, as the *test score* of a student in a common standardized exam like the SAT or ACT, and assume that more privileged students have access to testing; as Hyman (2016) notes, many well-qualified disadvantaged students do not have access to standardized tests and so cannot apply to schools that require them. On the other hand, as the University of California Standardized Testing Task Force (2020) and Bellafante (2020) posit, without testing it may be especially difficult to evaluate students from non-traditional backgrounds, as colleges instead rely on transcripts and recommendations from familiar (privileged) high schools. This aspect could be captured—as we do for our simulations—by considering the first $K - 1$ features as substantially more informative for group A ($\sigma_{Ak} < \sigma_{Bk}$), with a smaller informativeness discrepancy for the test score.

---

[8]We show, however, that these policies are insufficient for navigating the information-access tradeoff induced by the information requirements. Thus we focus on policies without affirmative action and, unless otherwise noted, the threshold on skill estimates is the same across all groups.

The model's focus differs from feature bias as traditionally understood, if a feature systematically under-values one group; e.g., weaker letters of recommendation for under-represented students. In our model, the school fully corrects for such bias (cancelling out μ_{gk}); in practice, schools interpret signals in context, for example benchmarking how many AP courses are offered by a student's school. In contrast, differential informativeness (function of σ_{gk}) and disparate access (γ_g) are harder to correct at admissions time. The former represents an information-theoretic limit to identifying the most qualified students, and the latter prevents some students from even applying. As we show, these effects cannot even be completely mitigated using affirmative action,[9] which is particularly insufficient in identifying qualified disadvantaged students.

Without loss of generality, we suppose that the features are less informative for group B than they are for group A. Specifically, let *unequal precisions* between groups mean $\sum_{k \in S} \sigma_{Ak}^{-2} > \sum_{k \in S} \sigma_{Bk}^{-2}$, and *equal precision* mean $\sum_{k \in S} \sigma_{Ak}^{-2} = \sum_{k \in S} \sigma_{Bk}^{-2}$. In settings with barriers, we assume that group A also has more access to the test, i.e., $\gamma_A \geq \gamma_B$.[10] Finally, the school is *selective* and has capacity $C < 1/2$. These assumptions are for exposition; our model's tractability allows us to solve analogously for the omitted cases.

3 Intuition: The role of differential informativeness

We begin our analysis in Section 3.1 by deriving how a Bayesian optimal school estimates the students' skill level. Then, we preview our main results, illustrating how the relationship between skill estimates and true skills of the applicant pool depends on the informativeness of features and the access barriers, with implications for how admissions differ by group.

3.1 School's optimal Bayesian estimation procedure

Our Bayesian school—with knowledge of the model parameters (feature noise means and variances)—observes each student's features and group membership and estimates their expected skill level, using properties of normal distributions. Repeating this process for all applicants induces the following distribution over the skill level estimates for each group.

Lemma 1 (Estimated skill). *Consider a school that uses feature set $S \subseteq \{1, \ldots, K\}$ for each applicant. Then, the perceived skill of an applicant in group $g \in \{A, B\}$ with feature values $\boldsymbol{\theta} = (\theta_k)_{k \in S}$ is:*

$$\tilde{q}(\boldsymbol{\theta}, g) = \frac{\mu \sigma^{-2} + \sum_{k \in S} (\theta_k - \mu_{gk}) \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2}}. \tag{1}$$

---

[9]In Section 4.3, we study our policies under the following definition of *affirmative action*: a constraint on the fraction of students from each group. This approach is common in the literature (Fang and Moro, 2011) and a proxy of the practices adopted by universities. However, as shown by recent lawsuits against Harvard (Hartocollis, 2019) and Yale (Hartocollis, 2020), the legal framework around affirmative action is ambiguous and restrictive. Explicit, predetermined racial quotas are generally illegal; conversely, University of Texas admits students using a high school-based quota system (The University of Texas, 2019).

[10]We further assume that, even in the presence of barriers, the market is *over-demanded* in the sense that the school can not admit all applicants, i.e., $C < (1 - \pi)\gamma_A + \pi\gamma_B$.
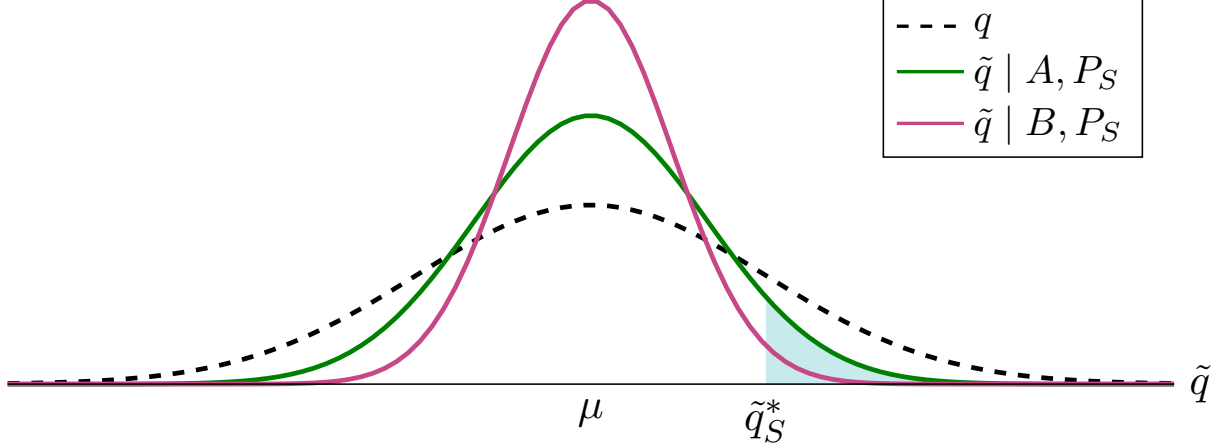
Figure 1: The distribution of skill estimates $\tilde{q}$ at an aggregate level for each group, as it depends on *informativeness* of the features. When application components are more precise for one group (group A, in green), the variance in the skill estimates of their group is *higher* – there is more signal for individuals to demonstrate that their skill is different than the mean. Then, more group A have high skill estimates above threshold $\tilde{q}_S^*$, and thus more are admitted. This effect occurs even though the true skill q distribution is identical across groups. If dropping the test causes such differential informativeness, then doing so may worsen both fairness and academic merit (estimated skill of admitted students). Figure 2 illustrates how the differential informativeness interacts with *disparate access*, due to which dropping test scores may improve all objectives.

Further, the skill level estimates for students in group g are normally distributed:

$$\tilde{q} \mid g, P_S \sim \mathcal{N}\left(\mu, \sigma^2 \left[\frac{\sum_{k \in S} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2}}\right]\right). \tag{2}$$

As Equation (1) shows, when the school estimates the skill level $\tilde{q}(\boldsymbol{\theta}, g)$ of an individual and knows the skill and feature noise distributions, it perfectly cancels out the mean bias terms μ_{gk} such that they do not affect estimation.[11] The school also re-weights each feature θ_k proportionately to the relative informativeness of this feature for group g: the less informative a feature is for a group (smaller *precision* σ_{gk}^{-2}), the less it contributes to estimates. Thus, due to informational differences in σ_{gk}^{-2} across groups, two students from different social groups with the same features $\boldsymbol{\theta}$ are evaluated differently. However, even in this idealized scenario, the school cannot fully correct for the variance terms σ_{gk}^2; two students with same skill q but in different groups have different skill estimates in expectation.

These individual estimation effects accumulate at the group level (Equation (2)) and drive our results on disparities. The school knows that $q \sim \mathcal{N}(\mu, \sigma^2)$ is identically distributed across social groups. However, as illustrated in Figure 1, the distribution $\tilde{q} \mid g, P_S$ for each group can differ across groups. For each group, the skill estimates are regularized toward the mean

---

[11](University of California Standardized Testing Task Force, 2020): "test scores are considered in the context of comprehensive review, which in effect re-scales the scores to help mitigate between-group differences."

(a) With test, assume no barriers (b) With test, inaccessible to some (c) Without test
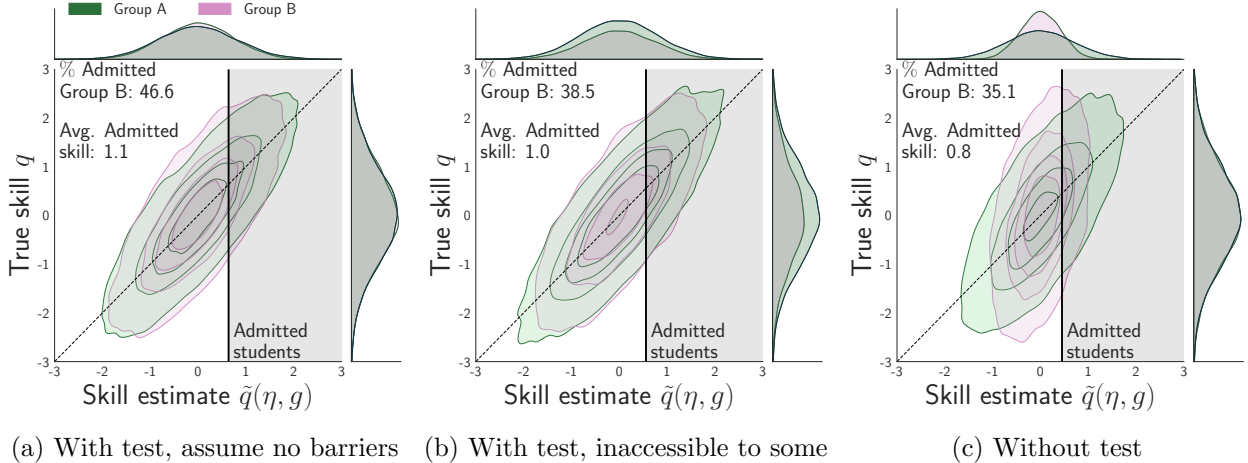
Figure 2: Skill vs estimate joint distribution for each group. Above and to the right of each joint distribution we plot the corresponding marginal distributions. The dashed diagonal lines correspond to perfect estimation. Full model parameters, as for all figures, can be found in Appendix A.5. Figure 2a represents a world without access barriers and when the features are approximately equally informative across groups. Figure 2b illustrates the consequences of requiring a test when group B (in pink) has access barriers: fewer can apply and so can be admitted. Figure 2c illustrates potential consequences of dropping the test: the school may be unable to distinguish among group B applicants, leading to worse estimates (rotated away from diagonal) and fewer admitted.

skill level μ. The regularization strength depends on the total precision $\sum_{k \in S} \sigma_{gk}^{-2}$: the larger the total precision for a group is (or the more informative its features are), the higher the variance in the estimated skills for that group is. In Figure 1, group A has larger total precision and for any value $\bar{q} > \mu$, there is a larger mass of students from group A than B with estimated skill higher than $\bar{q}$. When a college with capacity $C < \frac{1}{2}$ admits students with the highest skill estimates, more students in group A are admitted.

3.2 Intuition for impact of admissions policy

Before proceeding to our main results, we can now illustrate our main insight regarding the trade-off between informativeness and applicant pool size. In Figure 2, each sub-figure shows, for one scenario, the joint distribution between true skill q and the corresponding skill estimates $\tilde{q}$ for each group – along with the respective marginal distributions. Since both groups have identical true skill distributions, the joint distributions would *ideally* be identical for the two groups (and perfectly aligned along the diagonal); a policy that thresholds on estimated skill would then lead to group B composing a proportion π of the admitted class.

Consider the case where the potentially dropped feature (the "test score") is equally informative for both groups, whereas the remaining features are more informative for group A. Figure 2a illustrates the scenario when there are no access barriers to the test. Due to the differential informativeness induced by the other features, (slightly) more group A students are admitted: the college can better estimate their true skill, as illustrated by the group A joint distribution being closer

to the diagonal. Figure 2b illustrates the consequences of requiring test scores in the presence of unequal access levels ($\gamma_A = 1$ and $\gamma_B = \frac{2}{3}$). *Among those who apply*, the college can estimate true skill as well as it could in Figure 2a. However, fewer group B students can apply, as indicated by the smaller marginal count histogram, and so fewer are admitted. Figure 2c illustrates a scenario where the school removes the test score. Estimates for both groups are worse, as reflected in the joint distributions being further from the perfect estimation diagonal. However, skill estimates for group B students are especially degraded as their other features may be less informative, and so they make up a smaller proportion of the admitted class.

4   Theoretical results

In this section, we apply the insights from Section 3 on feature informativeness and skill estimation to our college admissions setting.

In Section 4.1, we focus solely on the effect of differences in informativeness, assuming no access barriers. We find that disparities arise with respect to all of our metrics of interest: academic merit, diversity, and individual fairness.

In Section 4.2, we compare two admissions policies: *with* and *without* a certain feature (e.g., test scores). When students have full and equal access to testing, we demonstrate how removing information might further decrease both fairness and academic merit under reasonable conditions. However, when students have different levels of access to the test, there is a trade-off between the barriers imposed by a test and the potentially valuable information a test may contain. Requiring the test can give schools more information on the students who do apply, but removing the test can increase the pool of skilled applicants. We characterize the school's optimal policy to include or exclude the test, depending on the relative sizes of these two effects.

Finally, in Section 4.3, we study the effect of affirmative action alongside the aforementioned policies. Affirmative action by definition improves diversity and individual fairness. However, it is insufficient in mitigating the differences in informativeness and, in the absence of barriers, its effect on the academic merit is heterogeneous: academic merit decreases further for group B and increases for group A.

4.1   Informational effects of fixed testing policies

In general, our fairness notions are not achievable – *even though we assume that both groups have the same true skill distributions.* For differential access barriers, this result immediately follows from the definitions: high skilled group B students who otherwise would be admitted can no longer apply. The effects of differential variance are more heterogeneous and are formalized in the result below.

Recall from Section 2 that $\tilde{q}_S^*$ denotes the admission threshold of the school under policy P_S. Let also Φ denote the CDF of the standard normal distribution $\mathcal{N}(0,1)$.

Proposition 1 (Metrics with a fixed policy). *Suppose that a selective school uses admissions policy P_S. Group fairness and individual fairness fail except for equal precision. Given unequal precisions:*

(i) *Diversity level: Group B students are under-represented, i.e., $\tau(P_S) < \pi$.*

   Furthermore, larger informativeness gap leads to decreased diversity: fix group B precision, $\sum_{k \in S} \sigma_{Bk}^{-2}$; then as group A precision increases, the diversity level $\tau(P_S)$ decreases.

(ii) *Individual fairness: High-skilled group B students are hard to target, i.e., $I(q; P_S) > 0$, if and only if*

$$q > \tilde{q}_S^* + \frac{\sigma^{-2}(\tilde{q}_S^* - \mu)}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}} \sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}.$$

   Increasing the informativeness gap increases the individual fairness gap for high-skilled students: fix group B precision, $\sum_{k \in S} \sigma_{Bk}^{-2}$; then as group A precision increases, $I(q; P_S)$ increases for $q > \mu + \sigma \Phi^{-1}(1 - C)$.

(iii) *Academic merit: The policy achieves worse academic merit for admitted students from group B.*

This result suggests that, although the school's Bayesian-optimal decision-making process can eliminate bias from skill estimates (see Section 3), the informativeness gap—as quantified via the difference in the total precision across groups—induces disparities in the admission outcomes even of ex-ante identical groups of students. As Figure 3 illustrates, and as we prove in Online Appendix C.3, with overall equal precision (the vertical line) both groups are admitted according to their population fractions (here, $1 - \pi = \pi = 0.5$); however, all fairness metrics degrade as the gap in informativeness between the two groups increases. Access barriers (even if limited to one group) have a similarly negative effect, albeit for a different reason: high-skilled students who otherwise would be admitted cannot even apply as they have not taken the test, cf. Hyman (2016).

The errors in estimation due to unequal precision affect the academic merit of each admitted group. Part (iii) establishes that under unequal precisions and no other disparities, students from group A admitted to selective colleges will be of higher true skill (on average) compared to the admitted students from group B, in contrast to existing theoretical results (cf., Faenza et al. (2020)). This discrepancy is due to the fact that the school fails to identify the high-skilled students from group B – part (ii) for individual fairness shows that high skilled students in group B are less likely to be admitted than they would if they were in group A.

We note that although the individual fairness gap is positive for all sufficiently high-skilled students, the magnitude of this gap varies. In fact, for students at the end of the right tail of the true skill distribution, the individual fairness gap starts to decrease, since – despite the noise – their estimates are high enough for admission. We prove this relationship in the following lemma.
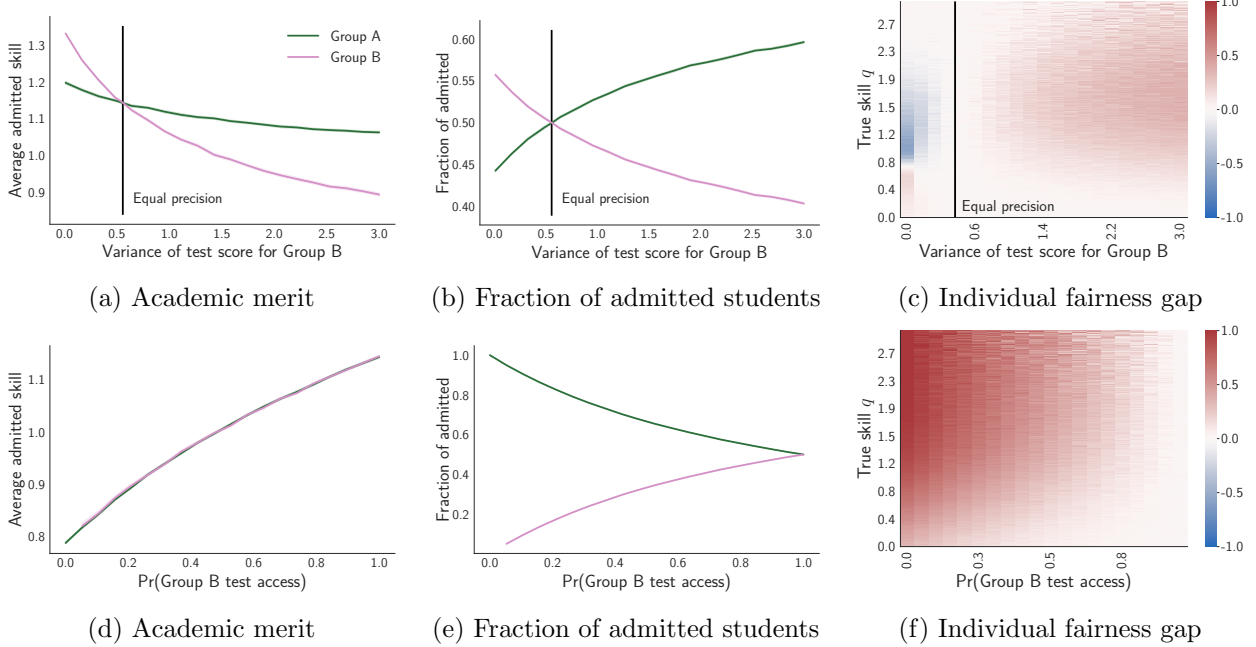
(a) Academic merit   (b) Fraction of admitted students   (c) Individual fairness gap

(d) Academic merit   (e) Fraction of admitted students   (f) Individual fairness gap

Figure 3: How the admitted students' academic merit, fraction of each group, and individual fairness gap change with group B test score variance and test access, respectively. With equal precision and no barriers, groups are treated equitably. As the feature variance or barriers increase for group B, both academic merit of admitted B students and fairness metrics worsen. We considered $\pi = 1 - \pi = 0.5$; the full parameter set can be found in Appendix A.5.

Lemma 2. *Consider policy P_S, and assume unequal precision. The individual fairness gap $I(q; P_S)$ is decreasing in q for $q > q_e$, where*

$$q_e \triangleq \tilde{q}_S^* + \sqrt{\frac{\sigma^{-4}(\mu - \tilde{q}_S^*)^2}{\sum_{k \in S} \sigma_{Ak}^{-2} \sum_{k \in S} \sigma_{Bk}^{-2}} + \frac{\ln\left(\sum_{k \in S} \sigma_{Ak}^{-2}\right) - \ln\left(\sum_{k \in S} \sigma_{Bk}^{-2}\right)}{\sum_{k \in S} \sigma_{Ak}^{-2} - \sum_{k \in S} \sigma_{Bk}^{-2}}}.$$

Furthermore, $\lim_{q \to \infty} I(q; P_S) = 0$.

Intuitively, a very high-skilled student has low probability to be incorrectly perceived as non-eligible for admission. This is because high $\tilde{q} > \tilde{q}_e$ overall leads to higher values of features $\boldsymbol{\theta}$ which in turn lead to a higher perceived skill $\tilde{q}$. Of course, the informational disadvantage of students in group B still has an effect and so the individual fairness gap remains positive.

These results on how a single policy performs as the model parameters change further hint at the difficulty in deciding whether to drop standardized testing. Doing so increases estimation variance (perhaps differentially, as Bellafante (2020) and University of California Standardized Testing Task Force (2020) posit), worsening all metrics, but also reduces access barriers, improving all metrics. These effects interact to induce the overall effect. Our next section formalizes this interaction.

13

4.2 Dropping test scores with and without barriers

In this subsection, we ask: *under what conditions would ignoring a feature benefit the school and the applicants?*. We study this question by comparing the test-free policy P_{SUB} to the test-based policy P_{FULL} in two different scenarios: Theorem 1 and Theorem 2 consider settings *with* and *without* *barriers*, respectively.

In Theorem 1 below, we study first the effect of dropping test scores on academic merit and diversity in the presence of barriers.

Theorem 1 (Dropping tests with barriers)**.** *Consider policies P_{FULL} and P_{SUB} and assume unequal precisions under P_{FULL}. In the presence of barriers, dropping test scores has the following implications:*

  (i) *Diversity level: Holding other parameters fixed, there exists threshold $\bar{\gamma}$ such that the diversity level improves under P_{SUB} if and only if $\gamma_B < \bar{\gamma}$.*

  (ii) *Academic merit: For each group g, holding other parameters fixed, there exists a threshold $\bar{\bar{\gamma}}_g$ such that academic merit of group g increases under P_{SUB} if and only if $\gamma_g < \bar{\bar{\gamma}}_g$.*

Perhaps surprisingly, Theorem 1 establishes that the academic merit of the admitted class may improve after dropping the test score. Similarly, diversity may deteriorate after dropping test scores. More specifically, Theorem 1 offers a threshold characterization, where the thresholds $\bar{\bar{\gamma}}_g$ and $\bar{\gamma}$ are functions of both the access levels of the two groups as well as the variance parameters, with and without the test. We provide the full characterization of these two quantities in Online Appendix C.5. We also include additional illustrations of the effects of dropping the test, with changes in the variance and access parameters; in particular, Figure 8 illustrates that the decision boundary in terms of the effect on diversity and the total and per-group academic merit can be non-linear.

At a high level, Theorem 1 implies that the decision to drop the test requirement is not just a matter of increasing access for the disadvantaged group. On the contrary, it depends on the complex interaction between the informational environment and the access levels of both groups. First, dropping test scores increases the applicant pool size but also affects its composition at different rates for each group. Second, the information loss incurred by dropping the test may not necessarily benefit students in group B. In particular, it is possible that the informational disadvantage faced by group B students may be exacerbated by the absence of test score information even if test scores are more noisy for group B than group A. In this case, the negative informational effect for group B may not be counterbalanced sufficiently by the increase in the group's pool size, especially when both groups face unequal, yet relatively proportional barriers.

In addition to the ambiguous impact that dropping test scores can have on the diversity of the admitted class and the academic merit of each group, the decision to drop the test introduces some additional implicit trade-offs. For example, as part (ii) in Theorem 1 indirectly implies and Figure 8

illustrates, only one group (e.g., group B) may be positively affected by this policy change,[12] even if the overall academic merit increases. Depending on the exact model parameters, this might be an inevitable consequence of dropping the test score. Nevertheless, it raises interesting and important fairness trade-offs for policy-makers.

Our next result studies the role of information loss in more depth, focusing on just the effect of the variance parameters in a setting without access barriers.

Theorem 2 (Dropping tests without barriers). *Consider policies P_{FULL} and P_{SUB}, and assume unequal precisions under P_{FULL}.*

(i) *Diversity level: Diversity level improves after dropping feature K, $\tau(P_{\text{SUB}}) > \tau(P_{\text{FULL}})$, if and only if*

$$\frac{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2} \left(\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2} \right)}{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2} \left(\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2} \right)} < \frac{\sigma_{AK}^{-2}}{\sigma_{BK}^{-2}}. \tag{3}$$

(ii) *Individual fairness: For each group g, there exist thresholds q_g such that the admission probability for students of skill q in group g decreases under P_{SUB} if and only if $q > q_g$. Further, there exists a threshold $\hat{q} \geq \max\{q_A, q_B\}$ such that the individual fairness gap increases for all $q > \hat{q}$, but may decrease otherwise.*

(iii) *Academic merit: Academic merit decreases for both groups $g \in \{A, B\}$, i.e.,*

$$\mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}] > \mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}].$$

In the absence of barriers, dropping a feature has mixed effects on the diversity level and individual fairness gap. However, it always worsens academic merit for both groups: without test scores, the school has access to fewer information signals and so skill estimates become noisier.

The exact effect on diversity depends on both the total precision of the remaining $K - 1$ features and how much the test precisions $\sigma_{A,K}^{-2}$, $\sigma_{B,K}^{-2}$ differ. More specifically, (3) is equivalent to the following condition

$$\frac{\frac{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}} < \frac{\frac{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}}, \tag{4}$$

which intuitively encodes how informativeness for each group changes after dropping the test. If (3) holds, then the diversity level improves as dropping the test narrows the relative informativeness gap between the two groups. However, if (3) does not hold (as University of California Standardized Testing Task Force (2020) attests), removing test scores exacerbates the informational disadvantage of students in group B. In that case, dropping the test decreases diversity.

Similarly, dropping the test may worsen individual fairness. As part (ii) shows, the admission probability of students with sufficiently high true skill, for either group, decreases after removing the

---

[12] As part (ii) in Proposition 2 shows, affirmative action has the same disproportionate effect across groups.

test. Furthermore, for sufficiently high-skilled students, the individual fairness gap increases after dropping test scores. This implication is independent of the actual effect on diversity; although the school may manage to improve diversity by dropping the test, the targeting of high-skilled students in both group becomes less effective, leaving high-skilled students in group B disproportionately affected compared to their same-skilled peers in group A.

Even without access barriers, the result establishes the importance of understanding features *other* than the test score – not just their biases (μ_{gk}, canceled out given full knowledge) but also their informativeness. More generally, our theoretical results illustrate that, even in a simple model, that the debate over dropping standardized testing cannot be had without the particulars of the context: whether one cares about overall academic merit of the admitted class or our fairness criteria, the effects depend on the relationships between access barriers, the information content of the test, and the information content of other application components. In Section 5, we will further demonstrate this result empirically and show how there exist real-world settings on either side of the divide – when dropping test score requirements would improve or worsen all our metrics.

4.3 Policies utilizing affirmative action

Schools have often an additional lever in their choice for admissions policy: whether or not to use affirmative action. In this section, we study the outcomes when schools can decide whether to require standardized testing *and* whether to use affirmative action. The term *affirmative action* describes admissions policies that partially base their decisions on applicants' membership in social groups with legally protected characteristics (e.g., race/ethnicity or gender), to support both equal opportunity and educational experiences diversity brings (Alon, 2015). These policies may thus use different admissions thresholds for different groups.

As a stylized model of affirmative action, we introduce a constraint on the diversity level $\tau(P)$ achieved by a policy P, i.e., consider admissions policies of the form P_S^τ, where $\tau \in (\tau(P_S), \pi]$ is the target diversity level set by the school. Thus, the school still optimizes for academic merit but under the additional constraint that a fraction τ of admitted students belong to group B. To do so, the common admission decision threshold is now replaced by two group-dependent thresholds, $\tilde{q}_{A,S}^*$ and $\tilde{q}_{B,S}^*$.[13] Note that $\tau(P_S^\tau) = \tau$; under affirmative action, diversity improves by definition, and group fairness holds when the target diversity level is set to $\tau = \pi$.[14] Affirmative action can be utilized on top of test-free or test-based policies. Whereas the testing policy determines the amount of information available in the estimation process, the affirmative action changes the selection process given information.

We find that although affirmative action increases diversity, it does not change the information

---

[13]In Proposition 2, the assumptions that $\gamma_A \geq \frac{2(1-\tau)C}{1-\pi}$ and $\gamma_B \geq \frac{2\tau C}{\pi}$ ensure that, even in the presence of barriers, the admission to the school is over-demanded (in the sense that the school cannot admit all applicants) and selective (meaning that the admission thresholds satisfy $\tilde{q}_{g,S}^* \geq \mu$).

[14]Proposition 2 focuses only on diversity levels $\tau \in (\tau(P_S), \pi]$. The lower bound is reasonable since $\tau(P_S)$ is the diversity level achieved by a school optimizing solely for academic merit (Theorem 1). The upper bound achieves group fairness. Note that higher levels $\tau > \pi$ could have also been considered with similar results; however, higher values of τ may be infeasible for certain values of C and $(1-\pi)$ therefore are omitted.

that schools have on students, and as a result the school still cannot identify high-skilled students in group B as well as it can identify group A students. We show that with unequal precision, affirmative action improves the individual fairness gap but does not eliminate it, as disparities in the identification of the highest-skilled students remain. It further increases the gap in academic merit across social groups. Affirmative action alone cannot address the fundamental issue caused by variance in the features. As a result, we consider this decision as orthogonal.

Proposition 2 (Affirmative action with a fixed testing policy). *Fix the target diversity level $\tau(P_S) < \tau \leq \pi$ and assume unequal precisions. Let also $\gamma_B \leq \gamma_A \leq 1$ such that $\gamma_A \geq \frac{2(1-\tau)C}{1-\pi}$, $\gamma_B \geq \frac{2\tau C}{\pi}$. Then,*

(i) Individual fairness: *In comparison to P_S, the individual fairness gap improves, i.e., $I(q; P_S^\tau) < I(q; P_S)$ for all q. However, group A students still have higher probability of admission than same-skilled group B students, i.e., $I(q; P_S) > 0$, if and only if*

$$
q > \frac{\left(\frac{\sum_{k\in S}\sigma_{Ak}^{-2}+\sigma^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}}\right)\tilde{q}_{A,S}^* - \left(\frac{\sum_{k\in S}\sigma_{Bk}^{-2}+\sigma^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}}\right)\tilde{q}_{B,S}^*}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} - \sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}} + \frac{\mu\sigma^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}}.
$$

Finally, there exist parameters such that $I(q; P_S^\tau) < 0 < I(q; P_S)$ for some q.

(ii) Academic merit: *Policy P_S^τ always achieves worse academic merit for admitted group B students than for group A students. Furthermore, in comparison to P_S, the academic merit of admitted students decreases for group B, while it increases for group A.*

We now study how test-free and test-based policies with affirmative actions compare in a setting with unequal barriers to test access. Recall that Theorem 1 shows (without affirmative action) that, conditional on the information environment, if there are substantial barriers to test access, removing the test requirement improves academic merit. The following theorem establishes the same result for a school using affirmative action. Let the function HR denote the hazard rate of the normal distribution Φ, $\mathrm{HR}(z) = \frac{\phi(z)}{1-\Phi(z)}$.

Proposition 3 (Dropping tests under affirmative action with barriers). *Fix group $g \in \{A, B\}$, variances σ_{gk}^2, and target diversity level τ. Let $\tau_A \triangleq 1 - \tau$ and $\tau_B \triangleq \tau$. Dropping the test score requirement improves the academic merit of admitted students from group g, i.e., $\mathbb{E}[q \mid Y = 1, g, P_{\mathrm{FULL}}^\tau] < \mathbb{E}[q \mid Y = 1, g, P_{\mathrm{SUB}}^\tau]$, if and only if $\gamma_g \leq \hat{\gamma}_g$, where*

$$
\hat{\gamma}_g = \frac{\tau_g C}{1 - \Phi\left(\mathrm{HR}^{-1}\left(\frac{\sqrt{\frac{\sum_{k\in\mathrm{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\mathrm{SUB}}\sigma_{gk}^{-2}}}}{\sqrt{\frac{\sum_{k\in\mathrm{FULL}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\mathrm{FULL}}\sigma_{gk}^{-2}}}}\mathrm{HR}(\Phi^{-1}(1 - \frac{\tau_g C}{\pi_g}))\right)\right)}. \tag{5}
$$

Fixing all other parameters, the threshold $\hat{\gamma}_g$ increases as test variance σ_{gK} for group g increases.

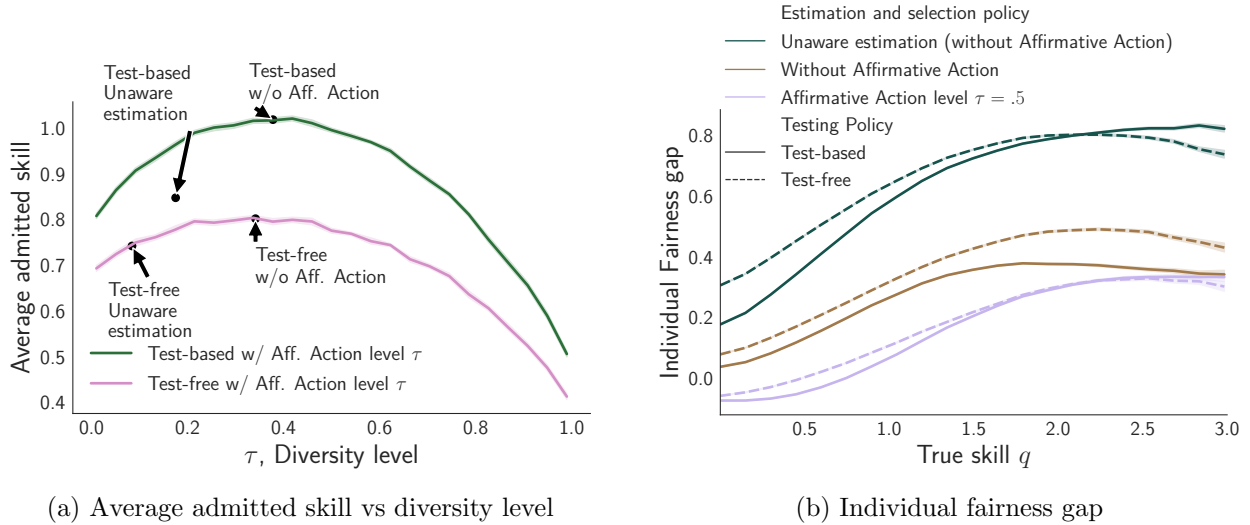(a) Average admitted skill vs diversity level

(b) Individual fairness gap

Figure 4: Performance of various policies – including with group-unaware estimation, in setting where features are more informative for group A, and with testing barriers for group B. Affirmative action in general improves both group diversity and individual fairness, while dropping the test score has an ambiguous impact. Group-unaware policies generally perform the worst on all metrics.

The threshold $\hat{\gamma}_g$ now depends only the characteristics of group g and τ, in contrast to Theorem 1, where the threshold depends on characteristics of both groups. The result further holds regardless of the economic inequality $\gamma_A - \gamma_B$ between the two groups; under affirmative action with a fixed diversity level, the school conducts the selection process for the two groups separately. Finally, as expected, if the test has a higher variance for a certain group, then it is more beneficial for that group to drop the test.

Comparing the policies. Figure 4 compares, for one parameter setting, our policies: with and without testing, and with and without affirmative action at various levels τ. In Figure 4a, the Pareto curves trace the trade-off between diversity and academic merit, for each testing policy. In this scenario, a constraint for group fairness (affirmative action at level $\tau = \pi = \frac{1}{2}$) does not substantially affect academic merit, while improving both group and individual fairness substantially. Furthermore, dropping tests has an ambiguous effect: it worsens diversity levels and academic merit, as well as the individual fairness gap in the case without affirmative action. However, it (slightly) improves the individual fairness gap with affirmative action.

Figure 4 also includes *group-unaware estimation* policies, that ignore the social group that a student belongs to; in this case, estimating student skill levels requires calculating the posterior from a mixture of Normal distributions. Ignoring group attributes is an oft-proposed but often problematic policy proposal to combat bias (Corbett-Davies and Goel, 2018). Perhaps unsurprisingly, group-unaware estimation policies perform most poorly. It worsens both the average academic merit of the admitted class and the diversity level, compared to the policy with group-aware estimation. It also leads to large individual fairness gaps, especially for high-skilled students. More details can be found in Online Appendix A.1.

Of course, all these effects are dependent on the parameter setting; we now turn to estimating these effects in data.

5 Calibrated simulations with college admissions data

Informational Case	Academic merit		Diversity Level	
	With test	Without test	With test	Without test
Low informativeness	3.25	3.23	16.6%	5.3%
High informativeness	3.49	3.51	25.0%	30.9%

Table 1: How academic merit and diversity level of admitted students change with and without requiring a test score, for two informational cases (how informative features other than the test score are). Whether dropping the test improves the objectives depends heavily on the informational environment: in one case, both objectives improve for all students after dropping test; in the other case, they both worsen. Note that group B students (from economically disadvantaged high schools) compose about 40% of the top 20% of students, but in all cases far fewer are accepted in our simulations.

Using application and transcript data from the University of Texas at Austin, we now demonstrate how an admissions committee could measure the trade-off in practice to better decide whether to drop their test scores requirement.

Data. Our data is from the Texas Higher Education Opportunity Project (THEOP), a public dataset of applications and transcripts for universities in Texas (Tienda and Sullivan, 2011). We focus on data from the University of Texas at Austin, for students who applied in 1992-1997.[15]

For each applicant, we observe their high school *class rank* (rounded to nearest decile), standardized *test score* (SAT, or ACT score translated to equivalent SAT score); we also observe demographic features (gender, ethnicity, citizenship status), characteristics of their high school (relative *economic privilege* rounded to nearest quartile, public/private, whether it's within Texas), and the major and college to which they're applying.[16] We further observe admissions decisions and, for accepted students, whether they enrolled. Finally, for those who enrolled, we observe rich transcript data: their GPA, number of credit hours, and major/college for each enrolled semester.

Simulation setup. We consider our applicant population as those who in reality *enrolled* to UT Austin (i.e., those for whom we observe college transcript data), and simulate a setting in which these applicants are further applying to a selective program, e.g., honors programs, scholarships, or college transfers. For each such individual, we use their cumulative college *college GPA* – not counting their first year – to represent their true skill. Then, as features, we use their high school *class rank, college* (e.g., Engineering, Business), and, in some cases, their standardized *test score* and college *first-year GPA*. To form the two groups, we take the upper (group *A*) and lower (group

---

[15]This period represents all applications from before the time Texas adopted the Top Ten Percent rule, in which all students at the top of their Texas public high school class were accepted regardless of other application components.

[16]The University of Texas admits students by college.

B) halves of the high schools' *economic privilege* index. We then simulate our model as follows. We train models using OLS to predict *college GPA* from the features, using all the students.[17] For each simulation, we sample an equal number of students from each group ($\pi = 0.5$) from the population, and assume that $\frac{2}{3}$ of group B students (and all group A students) have access to the test. Finally, we predict each student's *college GPA* using the features, either with or without *test score*, in one of two cases regarding how informative are features other than the test score:

Low informativeness *Class rank* and *college* are used as features.

High informativeness *Class rank*, *college*, and *first-year GPA* are used as features. The latter is far more predictive of final *college GPA* than are the other features.

Results. Table 1 shows the academic merit and diversity level of admitted students, for each informational case and with and without requiring the test. Results depend crucially on the informational environment. When the college has access to a high quality signal on all students – *first-year GPA* – dropping test scores increases both academic merit (for both groups) and diversity; it allows more students to apply, without incurring a substantial informational loss. In contrast, in the low informativeness case, without test scores the school must rely on students' high school ranks, which are especially uninformative for group B; then, such students are disproportionately rejected without test scores; *not* using test scores especially hurts already disadvantaged students as a group, even when $\frac{1}{3}$ of them are deemed to not have access to the test. In Appendix A.2, we show that these patterns are robust to the specific assumption on γ_B.

These findings underscore our theoretical results: the consequences of dropping test scores depends crucially on the information content of other signals, and the decision to do so should (and can) be made in a decision-driven manner.

Limitations. There are several major limitations to interpreting our analysis. First, we only observe application data for those students who were able to submit an SAT or ACT score, and so we rely on prior research (e.g., Hyman (2016)) to calibrate a population of students who *could have* applied in a world without a test score requirement. Second, GPA data is only available for students who are accepted to and who subsequently enroll at UT Austin. Thus, our data cannot help determine whether the test score is predictive of GPA success at admissions time.[18] Rather, our analysis should be interpreted as ranking *those already admitted*, such as for internal scholarships, transfers to other universities, or admissions to honors programs and colleges. Third, low-income and minority students face many challenges and barriers during their college education, and so their final GPA is itself not reflective of their true skill or academic merit (Engle and Tinto, 2008). Fourth, we do not observe several features available to admissions committees (such as recommendation letters), and so in practice test scores likely provide less marginal informativeness

---

[17]Results using a random forest model are qualitatively similar. We train a separate model for each group.

[18]In causal inference terms, being admitted is a *collider* between test scores and some unobserved true skill, which influences factors available to the admissions committee but not us as researchers, e.g., recommendation letters. This issue is a common barrier to measuring the predictive power of standardized testing (Weissman, 2020).

than in our simulation. Despite these limitations, however, we believe that our study demonstrates how an admissions committee with better data and potential to carry out experiments could make an informed decision on whether to drop testing requirements.

6   Discussion

Our work has policy implications beyond the formalization of the trade-off between information access and barriers in a testable framework. In Section 4.3, we showed that affirmative action (admitting the top students within each group) can improve diversity and individual fairness. However, it is insufficient in addressing the inequalities that arise due to differential informativeness and access barriers, as it neither helps schools identify the highest-performing students, nor does it increase the applicant pool size. Colleges must further invest in better signals and in expanding their applicant pools. In the setting where test scores are found to be highly effective for skill estimation *but* also impose large barriers, our work further suggests the value of another option for increasing fairness in admission: decreasing the access barriers. For example, several states have implemented policies to make the SAT and/or ACT mandatory for all public school students, while also reducing both financial and logistical barriers by paying the financial costs of test registration and offering the tests at more convenient times (Hyman, 2016).

Further, in reality individual schools do not make the decision to keep or drop testing requirements in isolation, but rather must react to the decisions that other schools make. When one school changes their own admission policy, and thus the pool of students they admit, other schools who are competing for the same pool of students may be affected. In Online Appendix B, we extend our model to study admissions decisions in settings where multiple schools compete for students and provide preliminary results. We show how students' preferences now affect the characteristics of the student body and that schools may have differing diversity levels even when using the same admission policy. We also begin to investigate the effect of one school's policy change on both its own students as well as those of the remaining schools.

Note that our theoretical results hold in a highly stylized setting where the school is Bayesian-optimal and knows the parameters of the model. Such an idealized scenario is, in practice, unattainable. We show that even under this idealized setting, inequalities arise – the school cannot correct for the differential informativeness in the features; our work thus presents an information-theoretic limit to how well schools can identify the most qualified students. Even if a school had full knowledge of each group's feature distributions (i.e., were able to perfectly evaluate students' skills in context), the school could not completely mitigate inequalities in admissions.

Another assumption is that all distributions are normal, which allows us to study the effect of variance in a transparent and tractable way. This assumption is not limiting: our results can be extended to a more general class of distributions such that group A's skill estimates are a mean-preserving spread (Blackwell, 1953) of group B's skill estimates, though analytic characterizations of the thresholds as we derive may not longer be possible.

We also started from the fundamental assumption that the two groups of students are equally

skilled; this approach is natural when the skill in our framework represents a student's 'potential.' Given that disparities arise even in this scenario with equal skill variables, we expect inequality – in terms of the individual fairness gap – to further worsen in a setting where one group is characterized by lower skill on average. Furthermore, our notion of barriers is restrictive; additional factors such as differential access to test preparation services (Park and Becks, 2015) and family support[19] (Espenshade and Radford, 2013; McDonough, 1997) may also constitute significant barriers for certain groups of students, though some of these may be captured in the noise *bias* term and corrected for by the school. However, our calibrated simulations in Section 5 show that our insights hold when these assumptions do not apply.

Overall, we believe that our work makes a novel modeling and conceptual contribution to the growing literature of fairness in decision-making systems. Our multi-feature take on the seminal model by Phelps (1972) naturally fits the study of fundamental questions related to fairness in operations, and can serve as a useful technical and conceptual framework to study emerging problems in fair algorithmic decision-making and public policy in education and beyond. More generally, our work suggests that the *design* of input features to machine learning tasks is an important challenge.

References

Atila Abdulkadiroğlu. 2005. College Admissions With Affirmative Action. *International Journal of Game Theory* 33, 4 (2005), 535–549.

Sigal Alon. 2015. *Race, Class, and Affirmative Action.* Russell Sage Foundation.

AJ Alvero, Sonia Giebel, Ben Gebre-Medhin, Anthony Lising Antonio, Mitchell L. Stevens, and Benjamin W. Domingue. 2021. Essay Content is Strongly Related to Household Income and SAT Scores: Evidence from 60,000 Undergraduate Applications. (2021).

Nick Anderson. 2020. Colleges are Ditching Required Admission Tests over Covid-19. Will They Ever Go Back? *The Washington Post* (June 2020). `www.washingtonpost.com/local/education/coronavirus-sat-act-admission/2020/06/15/18c406dc-acca-11ea-a9d9-a81c1a491c52_story.html`

Kenneth Arrow. 1971. The Theory of Discrimination. (1971).

Christopher Avery, Caroline Hoxby, Clement Jackson, Kaitlin Burek, Glenn Pope, and Mridula Raman. 2006. *Cost Should Be No Barrier: An Evaluation of the First Year of Harvard's Financial Aid Initiative.* Technical Report. National Bureau of Economic Research.

---

[19]For example, high socioeconomic status (SES) parents are more likely to be able to send their kids to private schools, hire a private tutor, help with homework, or move to neighborhoods with better public schools (Espenshade and Radford, 2013; McDonough, 1997).

Kathleen Aycock. 2021. Record Increase In Historically Underrepresented Graduate Applicants. `https://grad.berkeley.edu/news/announcements/record-high-increase-in-historically-underrepresented-graduate-applicants/`

Eduardo M Azevedo and Jacob D Leshno. 2016. A Supply and Demand Framework for Two-Sided Matching Markets. *Journal of Political Economy* 124, 5 (2016), 1235–1268.

Árpád Baricz. 2008. Mills' ratio: Monotonicity patterns and functional inequalities. *J. Math. Anal. Appl.* 340, 2 (2008), 1362–1370.

Gary S Becker. 1957. The Economics of Discrimination. *University of Chicago Press* (1957).

Ginia Bellafante. 2020. Should Ivy League Schools Randomly Select Students (At Least for a Little While)? *New York Times* (December 2020). `www.nytimes.com/2020/12/18/nyregion/ivy-league-admissions-lottery.html`

David Blackwell. 1953. Equivalent comparisons of experiments. *The annals of mathematical statistics* (1953), 265–272.

William Cai, Johann Gaebler, Nikhil Garg, and Sharad Goel. 2020. Fair Allocation Through Selective Information Acquisition. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (New York, NY, USA) *(AIES '20)*. 22–28.

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

Jimmy Chan and Erik Eyster. 2003. Does Banning Affirmative Action Lower College Student Quality? *American Economic Review* 93, 3 (2003), 858–872.

Stephen Coate and Glenn C Loury. 1993. Will Affirmative-Action Policies Eliminate Negative Stereotypes? *American Economic Review* (1993), 1220–1240.

Sam Corbett-Davies and Sharad Goel. 2018. The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning. *arXiv Preprint arXiv:1808.00023* (2018).

Bo Cowgill, Fabrizio Dell'Acqua, Samuel Deng, Daniel Hsu, Nakul Verma, and Augustin Chaintreau. 2020a. Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics. In *Proceedings of the 21st ACM Conference on Economics and Computation* (Virtual Event, Hungary) *(EC '20)*. 679–681.

Bo Cowgill, Fabrizio Dell'Acqua, Sam Deng, Daniel Hsu, Nakul Verma, and Augustin Chaintreau. 2020b. Replication Data (A) for 'Biased Programmers or Biased Data?': Individual Measures of Numeracy, Literacy and Problem Solving Skill – And Biographical Data – For a Representative Sample of 200K OECD Residents. `https://doi.org/10.7910/DVN/JAJ3CP`

Kuheli Dutt, Danielle L Pfaff, Ariel F Bernstein, Joseph S Dillard, and Caryn J Block. 2016. Gender Differences in Recommendation Letters for Postdoctoral Fellowships in Geoscience. *Nature Geoscience* 9, 11 (2016), 805–808.

Vitalii Emelianov, Nicolas Gast, Krishna P Gummadi, and Patrick Loiseau. 2020. On Fair Selection in the Presence of Implicit Variance. In *Proceedings of the 21st ACM Conference on Economics and Computation*. 649–675.

Jennifer Engle and Vincent Tinto. 2008. Moving beyond access: College success for low-income, first-generation students. *Pell Institute for the Study of Opportunity in Higher Education* (2008).

Dennis Epple, Richard Romano, and Holger Sieg. 2006. Admission, Tuition, and Financial Aid Policies in the Market for Higher Education. *Econometrica* 74, 4 (2006), 885–928.

Thomas J Espenshade and Alexandria Walton Radford. 2013. *No Longer Separate, Not Yet Equal: Race and Class in Elite College Admission and Campus Life*. Princeton University Press.

Yuri Faenza, Swati Gupta, and Xuan Zhang. 2020. Impact of Bias on School Admissions and Targeted Interventions. *arXiv Preprint arXiv:2004.10846* (2020).

Hanming Fang and Andrea Moro. 2011. Theories of Statistical Discrimination and Affirmative Action: A Survey. In *Handbook of Social Economics*. Vol. 1. Elsevier, 133–200.

Daniel Fershtman and Alessandro Pavan. 2020. Soft affirmative action and minority recruitment. *arXiv preprint arXiv:2004.14953* (2020).

Dean P Foster and Rakesh V Vohra. 1992. An Economic Argument for Affirmative Action. *Rationality and Society* 4, 2 (1992), 176–188.

Qiang Fu. 2006. A Theory of Affirmative Action in College Admissions. *Economic Inquiry* 44, 3 (2006), 420–428.

Nika Haghtalab, Nicole Immorlica, Brendan Lucier, and Jack Wang. 2020. Maximizing Welfare With Incentive-Aware Evaluation Mechanisms. In *29th International Joint Conference on Artificial Intelligence*.

Anemona Hartocollis. 2019. Harvard Does Not Discriminate Against Asian-Americans in Admissions, Judge Rules . *New York Times* (October 2019). `www.nytimes.com/2019/10/01/us/harvard-admissions-lawsuit.html`

Anemona Hartocollis. 2020. Justice Dept. Accuses Yale of Discrimination in Application Process . *New York Times* (August 2020). `www.nytimes.com/2020/08/13/us/yale-discrimination.html`

Virginia Hernanzi, Franck Malherbeti, and Michele Pellizzari. 2004. Take-Up of Welfare Benefits in OECD Countries. https://www.oecd-ilibrary.org/social-issues-migration-health/take-up-of-welfare-benefits-in-oecd-countries_525815265414

Lily Hu, Nicole Immorlica, and Jennifer Wortman Vaughan. 2019. The Disparate Effects of Strategic Manipulation. In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency*. 259–268.

Joshua Hyman. 2016. ACT for All: The Effect of Mandatory College Entrance Exams on Postsecondary Attainment and Choice. *Education Finance and Policy* 12 (05 2016), 1–69.

Nicole Immorlica, Katrina Ligett, and Juba Ziani. 2019. Access to Population-Level Signaling as a Source of Inequality. In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency*. 249–258.

Yuichiro Kamada and Fuhito Kojima. 2019. *Fair Matching Under Constraints: Theory and Applications*. Technical Report.

Sampath Kannan, Mingzi Niu, Aaron Roth, and Rakesh Vohra. 2021. Best vs. All: Equity and Accuracy of Standardized Test Score Reporting. *arXiv:2102.07809 [physics]* (Feb 2021). http://arxiv.org/abs/2102.07809 arXiv: 2102.07809.

Sampath Kannan, Aaron Roth, and Juba Ziani. 2019. Downstream Effects of Affirmative Action. In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency*. 240–248.

Jon Kleinberg and Sendhil Mullainathan. 2019. Simplicity Creates Inequity: Implications for Fairness, Stereotypes, and Interpretability. *arXiv:1809.04578 [cs, stat]* (Jun 2019). arXiv: 1809.04578.

Kevin Lang and Michael Manove. 2011. Education and Labor Market Discrimination. *American Economic Review* 101, 4 (2011), 1467–96.

Jonathan Levin. 2009. The Dynamics of Collective Reputation. *The BE Journal of Theoretical Economics* 9, 1 (2009).

Lydia T Liu, Ashia Wilson, Nika Haghtalab, Adam Tauman Kalai, Christian Borgs, and Jennifer Chayes. 2020. The Disparate Equilibria of Algorithmic Decision Making When Individuals Invest Rationally. In *Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency*. 381–391.

Zhi Liu and Nikhil Garg. 2021. Test-optional Policies: Overcoming Strategic Behavior and Informational Gaps. *arXiv preprint arXiv:2107.08922* (2021).

Patricia M McDonough. 1997. *Choosing colleges: How social class and schools structure opportunity*. Suny Press.

Hussein Mouzannar, Mesrob I Ohannessian, and Nathan Srebro. 2019. From Fair Decision Making to Social Equality. In *Proceedings of the Conference on Fairness, Accountability, and Transparency.* 359–368.

Giulia McDonnell Nieto del Rio. 2021. University of California Will No Longer Consider SAT and ACT Scores. *New York Times* (May 2021). `www.nytimes.com/2021/05/15/us/SAT-scores-uc-university-of-california.html`

Julie J Park and Ann H Becks. 2015. Who benefits from SAT prep?: An examination of high school context and race/ethnicity. *The Review of Higher Education* 39, 1 (2015), 1–23.

Edmund S Phelps. 1972. The Statistical Theory of Racism and Sexism. *American Economic Review* 62, 4 (1972), 659–661.

Richard Phelps. 2005. *Defending standardized testing.* Psychology Press.

Sean F Reardon. 2011. The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations. (2011).

Richard Reeves and Dimitrios Halikias. 2017. Race Gaps in SAT Scores Highlight Inequality and Hinder Upward Mobility. *Washington, DC: Brookings Institute* (2017).

Jeffrey Selingo. 2020. *Who Gets In and Why: A Year Inside College Admissions.* Scribner.

Emil Temnyalov. 2018. An Economic Theory of Differential Treatment. *SSRN Electronic Journal* (2018).

The University of Texas. 2019. Top 10 Percent Law. `news.utexas.edu/key-issues/top-10-percent-law/`

Marta Tienda and Teresa A Sullivan. 2011. *Texas Higher Education Opportunity Project.* Inter-university Consortium for Political and Social Research.

University of California Standardized Testing Task Force. 2020. Report of the UC Academic Council Standardized Testing Task Force. `https://senate.universityofcalifornia.edu/_files/underreview/sttf-report.pdf`

Michael B Weissman. 2020. Do GRE scores help predict getting a physics Ph.D.? A comment on a paper by Miller et al. *Science Advances* 6, 23 (2020).

Rebecca Zwick. 2002. *Fair Game?: The Use of Standardized Admissions Tests in Higher Education.* Psychology Press.

A Supplementary results and figures

A.1 Group-unaware estimation

In the main text, we primarily consider a "group-aware" estimation procedure, in which the school uses students' group membership in its estimation procedure (and thus is able to plug in group-specific noise biases and variances). We now briefly discuss "unaware" estimation when it cannot do so. Ignoring group attributes is an oft-proposed but often problematic policy proposal to combat bias in machine learning tasks (Corbett-Davies and Goel, 2018), and so we evaluate its consequences.

Ignoring group membership complicates the skill estimation challenge. When the feature distributions differ across groups but the school cannot observe the group of a student, the resulting estimated skill distribution is a mixture of normal distributions. The mixture weights depend on the noise means and variances of each group g. In contrast to the group-aware case, where the school manages to correct for the feature noise biases (but not variance), the biases now play an important rule in each feature's implications.

We derive this distribution below. However, we primarily study the effects through simulation in Figure 4.

Unaware estimation derivation. Conditional on the true skill level q, the features are still distributed according to a group-specific Normal distribution:

$$\theta_k | q, g \sim N(q + \mu_{gk}, \sigma_{gk}^2) \qquad \forall k = 1 \ldots K$$

But under group-unaware estimation, the school does not know or cannot use g, so the posterior is now a mixture of Normal distributions. Specifically, let $f(q \mid \boldsymbol{\theta})$ denote the pdf of the posterior distribution, $q \mid \boldsymbol{\theta}$; similarly, we use the notation $f(\boldsymbol{\theta})$ and $f(q \mid \boldsymbol{\theta}, g)$. Thus,

$$
\begin{aligned}
f(q | \boldsymbol{\theta}) &= \sum_{g \in \{A,B\}} f(q | \boldsymbol{\theta}, g) \, \mathbb{P}(g | \boldsymbol{\theta}) \\
&= \sum_{g \in \{A,B\}} f(q | \boldsymbol{\theta}, g) \left[\frac{f(\boldsymbol{\theta} | g) \, \mathbb{P}(g)}{f(\boldsymbol{\theta})} \right] \\
&= \sum_{g \in \{A,B\}} w(\boldsymbol{\theta}, g) f(q | \boldsymbol{\theta}, g), \qquad\qquad w(\boldsymbol{\theta}, g) \triangleq \left[\frac{f(\boldsymbol{\theta} | g) \, \mathbb{P}(g)}{f(\boldsymbol{\theta})} \right].
\end{aligned}
$$

Then, the posterior $q | \boldsymbol{\theta}$ is distributed as a mixture of Normal distributions, where each Normal is as in the group-aware case:

$$q | \boldsymbol{\theta} \sim \sum_{g \in \{A,B\}} w(\boldsymbol{\theta}, g) N\left(\tilde{q}(\boldsymbol{\theta}, g), \tilde{\sigma}^2(\boldsymbol{\theta}, g) \right)$$

For the weights, we find that

$$w(\boldsymbol{\theta}, g) \triangleq \frac{f(\boldsymbol{\theta} \mid g)\, \mathbb{P}(g)}{f(\boldsymbol{\theta})} = \frac{\int_{\infty}^{\infty} \Pi_k f(\theta_k \mid g, q)\, \mathrm{d}F(q) \cdot \mathbb{P}(g)}{f(\boldsymbol{\theta})}$$

and for K features,

$$\int_{\infty}^{\infty} \Pi_k f(\theta_k \mid g, q)\, \mathrm{d}F(q) = \frac{e^{\left(-\frac{\sum_{k=1}^{K}\left[\left(\mu+\mu_{gk}-\theta_k\right)^2 \sigma^{-2}\sigma_{gk}^{-2}\right] + \sum_{k\neq\ell}\left[\left((\mu_{\ell g}-\theta_\ell)-(\mu_{gk}-\theta_k)\right)^2\sigma_{\ell g}^{-2}\sigma_{gk}^{-2}\right]}{2\left(\sigma^{-2}+\sum_{k=1}^{K}\sigma_{gk}^{-2}\right)}\right)}}{2\left(1-\pi\right)^{K/2}\sigma\left(\Pi_k \sigma_{gk}\right)\sqrt{\sigma^{-2}+\sum_{k=1}^{K}\sigma_{gk}^{-2}}} \tag{6}$$

Thus, we have

$$w(\boldsymbol{\theta}, g) \triangleq \frac{f(\boldsymbol{\theta} \mid g)\, \mathbb{P}(g)}{f(\boldsymbol{\theta})} = \frac{\int_{\infty}^{\infty} \Pi_k f(\theta_k \mid g, q)\, \mathrm{d}F(q)\, \mathbb{P}(g)}{f(\boldsymbol{\theta})}$$

$$\propto \frac{\mathbb{P}(g)\exp\left\{\left(-\frac{\sum_{k=1}^{K}\left[\left(\mu+\mu_{gk}-\theta_k\right)^2\sigma^{-2}\sigma_{gk}^{-2}\right]+\sum_{k\neq\ell}\left[\left((\mu_{\ell g}-\theta_\ell)-(\mu_{gk}-\theta_k)\right)^2\sigma_{\ell g}^{-2}\sigma_{gk}^{-2}\right]}{2\left(\sigma^{-2}+\sum_{k=1}^{K}\sigma_{gk}^{-2}\right)}\right)\right\}}{[\Pi_k \sigma_{gk}]\sqrt{\sigma^{-2}+\sum_{k=1}^{K}\sigma_{gk}^{-2}}}$$

Derivation for equation (6). We explicitly show the algebra for $K = 1$ and $K = 2$ features, and the pattern continues for K features.

For one feature:

$$w(\theta_1, g) \triangleq \frac{f(\theta_1 \mid g)\, \mathbb{P}(g)}{f(\theta_1)} = \frac{\int_{\infty}^{\infty} f(\theta_1 \mid g, q)\, \mathrm{d}F(q)\, \mathbb{P}(g)}{f(\theta_1)}$$

$$= \frac{\frac{1}{\sqrt{2(1-\pi)(\sigma^2+\sigma_{g1}^2)}}\exp\left[-\frac{(\mu+\mu_{g1}-\theta_1)^2}{2(\sigma^2+\sigma_{g1}^2)}\right]\mathbb{P}(g)}{f(\theta_1)} = \frac{\frac{1}{\sqrt{(\sigma^2+\sigma_{g1}^2)}}\exp\left[-\frac{(\mu+\mu_{g1}-\theta_1)^2}{2(\sigma^2+\sigma_{g1}^2)}\right]\mathbb{P}(g)}{\sum_g\left[\frac{1}{\sqrt{(\sigma^2+\sigma_{g1}^2)}}\exp\left[-\frac{(\mu+\mu_{g1}-\theta_1)^2}{2(\sigma^2+\sigma_{g1}^2)}\right]\mathbb{P}(g)\right]}$$

For two features θ_1, θ_2:

$$w(\boldsymbol{\theta}, g) \triangleq \frac{f(\boldsymbol{\theta} \mid g)\, \mathbb{P}(g)}{f(\boldsymbol{\theta})} = \frac{\int_{\infty}^{\infty} \Pi_k f(\theta_k \mid g, q)\, \mathrm{d}F(q)\, \mathbb{P}(g)}{f(\boldsymbol{\theta})}$$

$$\int_{\infty}^{\infty} \Pi_k f(\theta_k \mid g, q)\, \mathrm{d}F(q) = \frac{e^{\left(-\frac{\left(\left(\mu_{g1}-\theta_1\right)-\left(\mu_{g2}-\theta_2\right)\right)^2\sigma_{g1}^{-2}\sigma_{g2}^{-2}+\left(\mu+\mu_{g2}-\theta_2\right)^2\sigma^{-2}\sigma_{g2}^{-2}+\left(\mu+\mu_{g1}-\theta_1\right)^2\sigma^{-2}\sigma_{g1}^{-2}}{2\left(\sigma^{-2}+\sigma_{g1}^{-2}+\sigma_{g2}^{-2}\right)}\right)}}{2\left(1-\pi\right)\sigma\sigma_{g1}\sigma_{g2}\sqrt{\sigma^{-2}+\sigma_{g1}^{-2}+\sigma_{g2}^{-2}}}$$

(a) With SAT, assume no barriers   (b) With SAT and barriers   (c) Without SAT

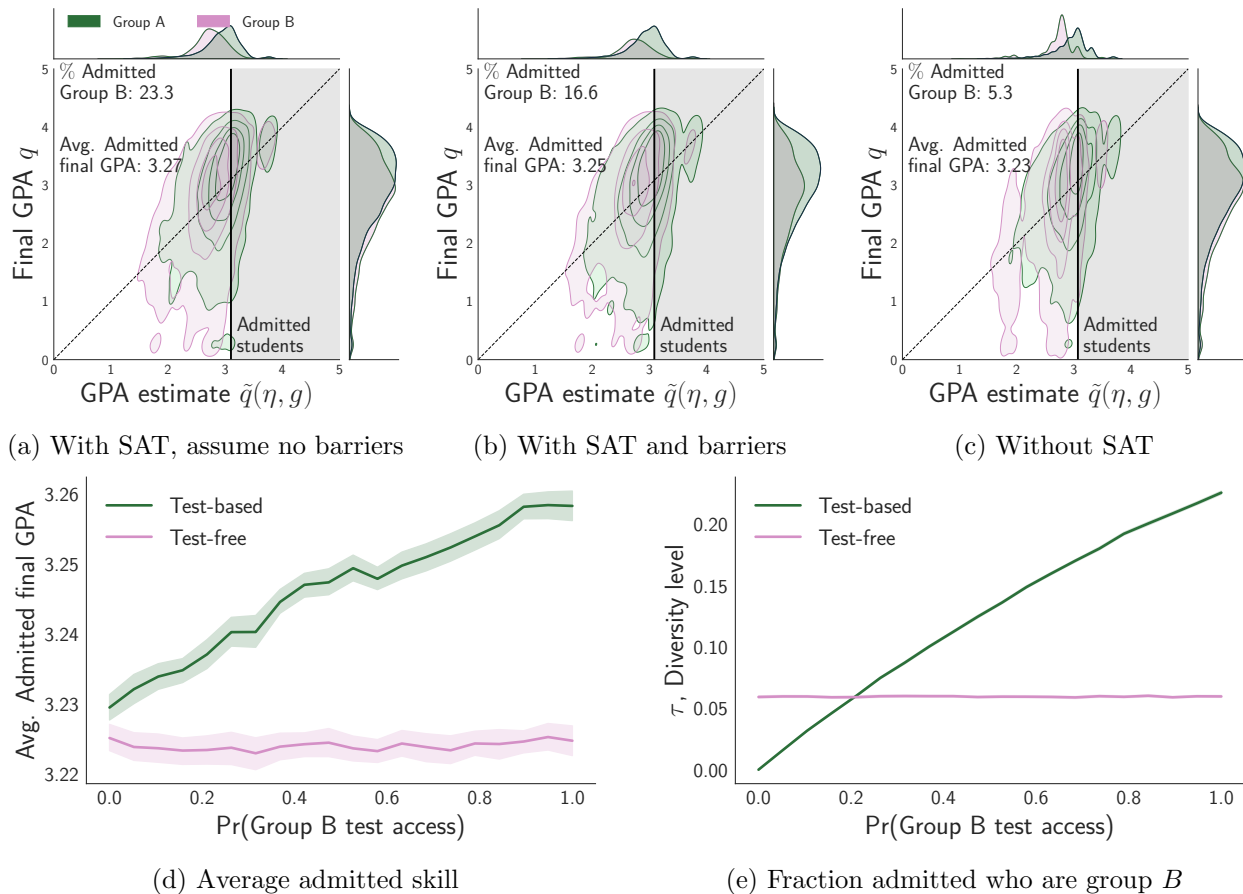(d) Average admitted skill   (e) Fraction admitted who are group B

Figure 5: Additional information for the *Low Information* case, when the standardized test scores provide much more information to the school than do the other features. In this setting, keeping test scores is best for both objectives, except for extremely low γ_B.

A.2   UT Austin data simulations supplementary plots

Figures 5 and 6 show additional information for the Low and High informativeness cases, respectively. The results demonstrate that the values reported in Table 1 in the main text are robust to our assumption on the fraction of group B students γ_B that have access to the test score: in these cases, the informativeness differences between the cases determine whether dropping test scores benefits the dual-objectives.

A.3   Second data setting: OECD study

We now repeat our calibrated simulation analysis, using an alternate data source, to analyze the robustness of our methods and insights.

Data. We now use a rich cross-national dataset consisting of demographic information and the results of a standardized test for about 200,000 people, collected starting in 2008 by the Organisation for Economic Co-operation and Development (OECD) as part of the Programme for the

(a) With SAT, assume no barriers   (b) With SAT and barriers   (c) Without SAT

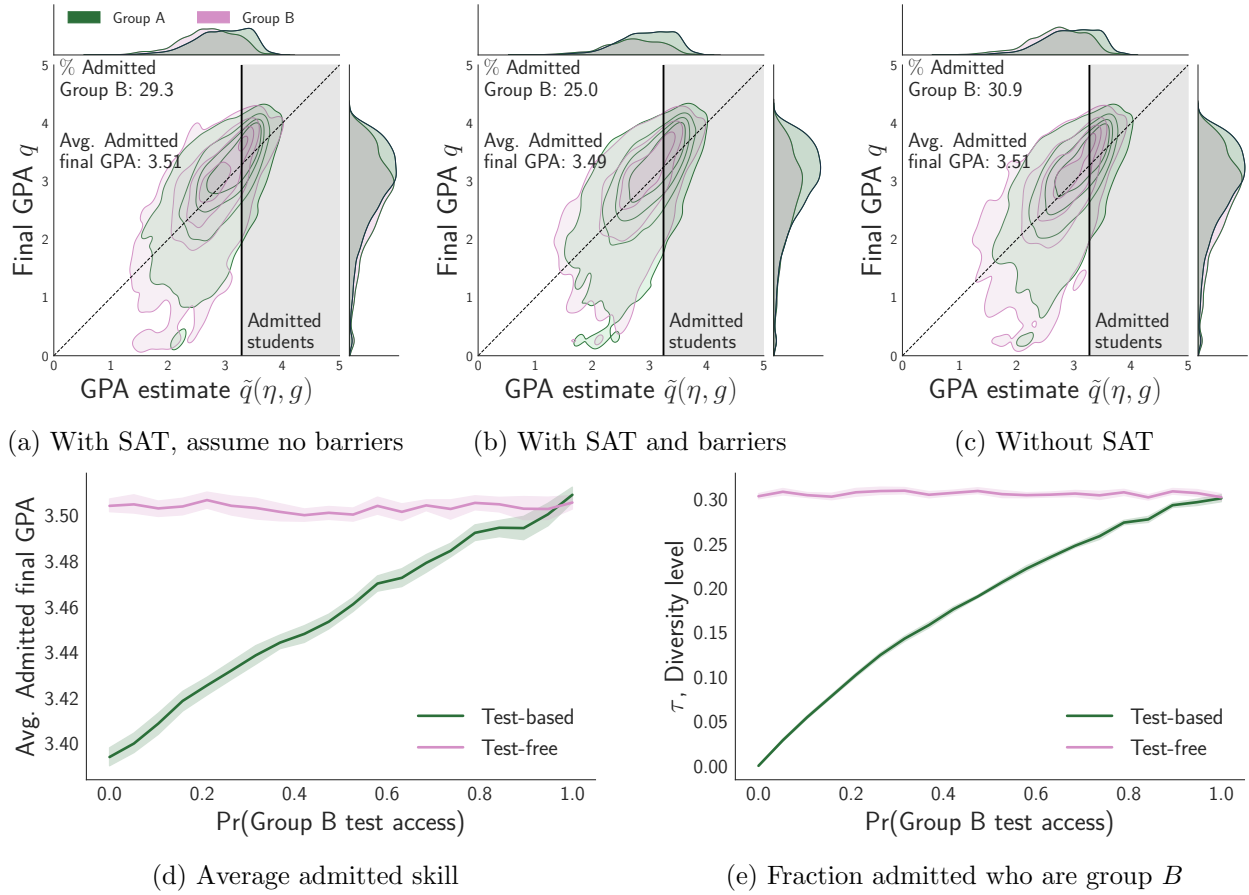(d) Average admitted skill   (e) Fraction admitted who are group B

Figure 6: Additional information for the *High Information* case, when the standardized test scores do not provide much more information to the school than do the other features. In this setting, dropping test scores is best for both objectives, across the range of γ_B.

International Assessment of Adult Competencies.[20]

(a) With test, assume no barriers   (b) With test, inaccessible to some   (c) Without test

(d) Average admitted skill   (e) Fraction admitted who are group B
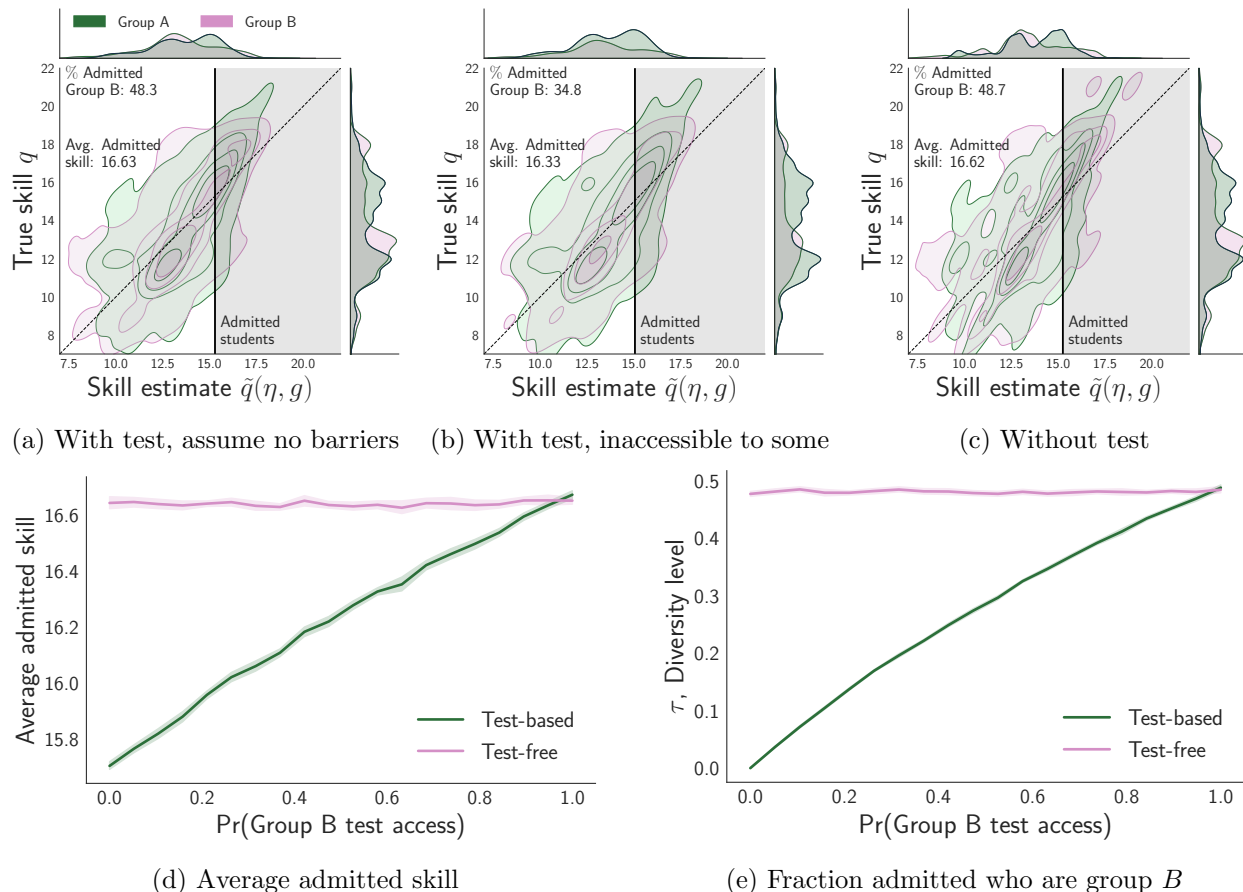
Figure 7: Calibrated simulation results for setting described in Appendix A.3.

Simulation setup. For each individual, we observe their *birth region*, *total years of education*, *years of education needed for their current job*, *normalized monthly income*, and their *test score*.[21] We use *birth region* as the person's group membership, limiting to two regions: *North America and Western Europe* (group A), and *Central and Eastern Europe* (group B). As we do not observe the "true" skill level of each student, we use *years of education* as their skill level. Then, *years of education needed for current job* and *test scores* make up our two features. Finally, we threshold[22] on *normalized monthly income* to determine whether the individual faces barriers to taking the test. Group A students on average have much higher incomes in the data; 73% of them have access to the test, versus 23% of group B. Note that this thresholding approach allows for general correlations between the features and ability to afford the test, unlike in our theoretical setting. We then simulate our model, as in Section 5.

Results. Figure 7 contains results. As in the *High Informativeness* setting with the THEOP data,

---

[20]We access a cross-country merged dataset (Cowgill et al., 2020b) compiled by Cowgill et al. (2020a).

[21]Test score is derived as described in the release notes of Cowgill et al. (2020b).
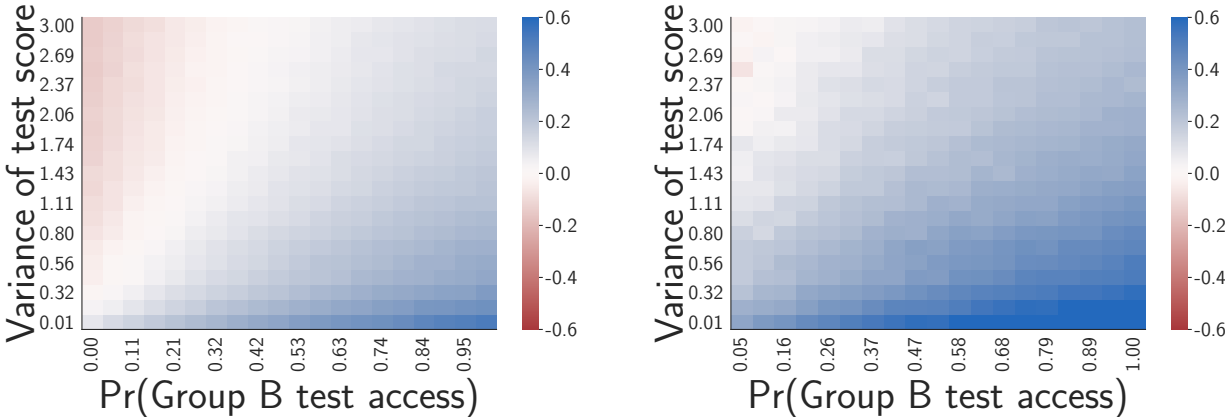
[22]We arbitrarily choose a threshold of $2000, but the result is robust to this choice.

we find that in this setting dropping the "test" feature substantially improves both diversity and academic merit, in comparison to requiring an inaccessible test.
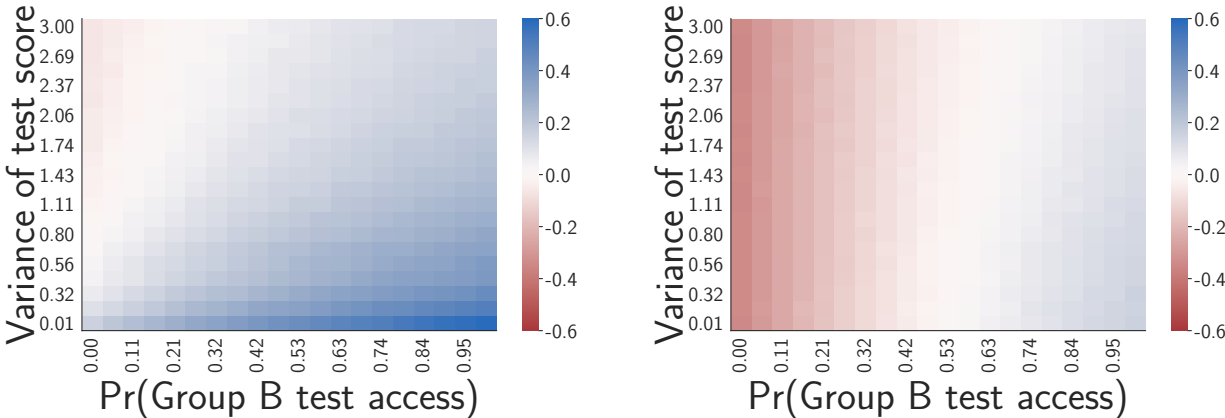
(a) Difference in Average admitted skill

(b) Difference in Fraction group B admitted

(c) Difference in Average admitted skill group A

(d) Difference in Average admitted skill group B

Figure 8: Difference between test-based and test-free policies with respect to various objective functions. The more negative (red) the difference, the more that dropping the test *improves* that metric compared to test-based policies. Simulation is with budgets case, using parameters as given in Appendix A.5. The plot reads as follows: in Figure 8a, a difference of 0.6 means that the average academic merit with a test-based policy is 0.6 higher than that with a test-free policy.

A.4 Supplemental simulation figures

Figure 8 supplements the results in Theorem 1 and Proposition 3, regarding the thresholds at which academic merit and diversity improve after dropping the test. In particular, they illustrate that for high enough test score variance or high enough barriers, dropping the test score improves the objectives.

A.5 Simulation parameters

Figure 2. $C = 0.2, \pi = 0.5, q, \theta_{A0}, \theta_{A1} \sim N(0, 1), \theta_{B0} \sim N(-4, 5), \theta_{B1} \sim N(-4, 1), \gamma_A = 1, \gamma_B = \frac{2}{3}$.

Figure 3. Same as Figure 2, except with $\theta_{B1} \sim N(-4, \sigma_{B1}^2)$, where $\sigma_{B1}^2 \in (0, 5)$, and $\gamma_B = 1$.

Figure 4. Same as Figure 2.

Figure 8. Same as Figure 2, except with test score precision varying together for both groups $\sigma_{A1}^2 = \sigma_{B1}^2 \in (0, 3)$, and group B test access varying, $\gamma_B \in (0, 1)$.

B   Multiple Schools

In reality, individual schools do not make the decision to keep or drop testing requirements in isolation, but rather must react to the decisions that other schools make. When one school changes their own admission policy, and thus the pool of students they admit, other schools who are competing for the same pool of students may be affected. In this section, we describe how to extend the model described in Section 2 to study admissions decisions in settings where multiple schools compete for students. We provide some preliminary results to suggest possible directions to study in this setting.

When there are multiple schools, a student who is admitted to more than one school can choose which school they prefer. Thus the students' preferences create a discrepancy between the set of students *admitted* to a school and the set of students who choose to *attend* a school. We investigate the impact of interventions on the set of students who ultimately attend the school, in terms of both academic merit and fairness.

We show that this competition between schools creates additional diversity concerns not present in the single-school case. Furthermore, when one school makes a change to their admissions policy, this change has downstream effects on other schools.

B.1   Multiple school model

The model is as defined in Section 2, with the following changes.

Schools. A finite set of schools $A_1, \ldots, A_N$ separately admit students. Each school A_j has capacity C_j for the mass of students who *attend* the school. The market is over-demanded: $\sum_{j=1}^{N} C_j < 1$.

Each school can choose their own estimation and admissions policy. Let $P(A_j)$ denote the admissions policy of school A_j. Let $\boldsymbol{P}$ denote the vector of policies for schools $A_1, \ldots, A_N$. Note that schools with different estimation policies may assign different perceived skills to a given student, and so the ranking of students according to perceived skill may not be consistent across schools.

Students' preferences. Students have common preferences over schools, $A_1 \succ \ldots \succ A_N$. Students prefer any school over their outside option A_o (i.e., $A_N \succ A_o$) and so they apply to all schools, unless they are prevented from doing so by testing barriers. We make the assumption of common preferences for simplicity, although the model can easily be extended so that students have heterogeneous preferences.

Admission outcomes. As in the single school case, each school chooses a selection policy, inducing an admissions threshold, and admits all students with perceived skill greater than the threshold.

Then, each student attends the school that they most prefer, among the set of schools to which they were admitted. If the student was not admitted to any school, then they choose their outside option A_o. These outcomes correspond to a matching of students to schools. Extending the notation from Section 2, for a given student let $Y_j \in \{0, 1\}$ denote whether the student attends school A_j.[23]

There exists a unique set of thresholds, which depends on the *entire* vector of policies chosen by the schools, such that each school fills its capacity with its most preferred students not accepted to a higher-ranked school.[24]

B.2   Benchmark: Identical school policies

First we consider the case when all schools employ the same admission policy P_S. In this case, we can characterize the stable matching of the system as follows.

Proposition 4 (Equilibrium with identical policies.). *Suppose that students have a common ranking of the schools and that schools all see the same feature set $S \subseteq \{1, \ldots, K\}$ and use a group-aware admissions policy where they admit on perceived quality $\tilde{q}$.*

Then, a stable matching is characterized by a set of cutoffs $\tilde{q}_1^* > \ldots > \tilde{q}_j^* > \ldots > \tilde{q}_N^*$ such that school A_1 is matched to students with perceived quality $\tilde{q}(\boldsymbol{\theta}, g, P_S) \geq \tilde{q}_1^*$ and for $j \geq 2$, school A_j is matched to students with perceived skill $\tilde{q}_j^* \leq \tilde{q}(\boldsymbol{\theta}, g, P_S) < \tilde{q}_{j-1}^*$.

The cutoffs can be found by solving the equations

$$C_j = \pi \int_{q_{j-1}^*}^{q_j^*} f_{\tilde{q}|A,P_S}(\tilde{q})d\tilde{q} + (1-\pi) \int_{q_{j-1}^*}^{q_j^*} f_{\tilde{q}|B,P_S}(\tilde{q})d\tilde{q}.$$

Proof. The result follows directly from an application of (Azevedo and Leshno, 2016). □

Recall that in the single school case with a selective school, Proposition 1 shows that disadvantaged students are under-represented unless equal precision holds. In general, this result does not hold for all schools in the multiple school setting.

The following proposition shows that whether a group is under-represented at a given school depends on the capacity constraints of other schools. We define three classes of schools, *top-tier*, *mid-tier*, and *low-tier*, which depend on both the students' ranking of the schools and the schools' capacities. We then show that whether disadvantaged students are under-represented at a school depends on the tier of the school.

Proposition 5 (Identical school policies). *Assume unequal precision and that all schools are selective with capacity at most 0.5. Suppose that schools use admissions policy P_S for some set feature set S. For a school A_j, define $D_j = \sum_{j' \leq j} C_{j'}$ to be the cumulative capacity of school A_j and all schools $A_1, \ldots, A_{j-1}$ that students prefer over A_j.*

---

[23]In Section 2, we defined $Y = 1$ for students who are admitted, but this set of students is equivalent to those who attend in the single school case.

[24]Our model is a special case of the two-sided market framework of Azevedo and Leshno (2016), with a common ranking of schools. The set of thresholds corresponds to a stable matching, unique under mild conditions. See Proposition 4.
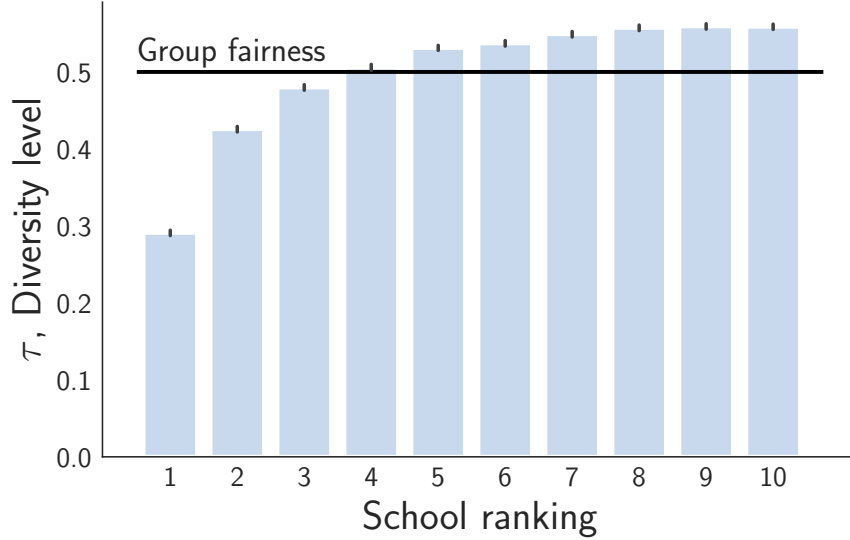
Figure 9: Diversity levels for each school in the multiple school setting, when all schools use the benchmark group-aware, test-based policy. Ten schools, each with capacity $C = 0.05$. Student parameters as in Figure 4.

Then there exists two capacity cutoffs $0 < X_1 < X_2 < 1$ such that:

- *Top-tier schools: When $1 - D_j > X_2$, group B students are under-represented.*

- *Mid-tier schools: When $X_1 < 1 - D_{j-1} < X_2$ and $X_1 < 1 - D_j < X_2$, group B students are not under-represented.*

- *Low-tier schools: When $1 - D_{j-1} < X_1$, group B students are under-represented.*

Proof. Consider the distributions of $\tilde{q} \mid A, P_S$ and $\tilde{q} \mid B, P_S$ and the respective pdfs $f_{\tilde{q}\mid A,P_S}, f_{\tilde{q}\mid B,P_S}$. Both distributions are normally distributed with the same and potentially difference variance. Assuming $\sum_{k \in S} \sigma_{Ak}^{-2} \neq \sum_{k \in S} \sigma_{Bk}^{-2}$, then $\mathrm{Var}(\tilde{q} \mid A) \neq \mathrm{Var}(\tilde{q} \mid B, P_S)$ and $f_{\tilde{q}\mid A,P_S}$ and $f_{\tilde{q}\mid B,P_S}$ cross at exactly two points q that solve the following equation:

$$2\frac{\mathrm{Var}(\tilde{q} \mid A, P_S) \cdot \mathrm{Var}(\tilde{q} \mid B, P_S)}{\mathrm{Var}(\tilde{q} \mid B, P_S) - \mathrm{Var}(\tilde{q} \mid A, P_S)} \log\left(\frac{\sqrt{\mathrm{Var}(\tilde{q} \mid B, P_S)}}{\sqrt{\mathrm{Var}(\tilde{q} \mid A, P_S)}}\right) = (q - \mu)^2.$$

Denote the solutions q_1, q_2, with $q_1 < q_2$.

Note that for any interval $[c, d]$, the mass of students in group A with $\tilde{q}(\boldsymbol{\theta}, A) \in [c, d]$ is

$$\pi \int_c^d f_{\tilde{q}\mid A,P_S}(\tilde{q})d\tilde{q}$$

and the mass of students in group B with $\tilde{q}(\boldsymbol{\theta}, B) \in [c, d]$ is

$$\int_c^d f_{\tilde{q}\mid B,P_S}(\tilde{q})d\tilde{q}.$$

35

Now define the capacity cutoffs

$$X_1 = \pi \int_{q_1}^{\infty} f_{\tilde{q}|A,P_S}(\tilde{q})d\tilde{q} + (1-\pi) \int_{q_1}^{\infty} f_{\tilde{q}|B,P_S}(\tilde{q})d\tilde{q}$$

$$X_2 = \pi \int_{q_2}^{\infty} f_{\tilde{q}|A,P_S}(\tilde{q})d\tilde{q} + (1-\pi) \int_{q_2}^{\infty} f_{\tilde{q}|B,P_S}(\tilde{q})d\tilde{q}.$$

If $\sum_{k=1}^{K} \sigma_{Ak}^{-2} > \sum_{k=1}^{K} \sigma_{Bk}^{-2}$, then $f_{\tilde{q}|A,P_S}(\tilde{q}) < f_{\tilde{q}|B,P_S}(\tilde{q})$ for $\tilde{q} \in [q_1, q_2]$ and $f_{\tilde{q}|A,P_S}(\tilde{q}) > f_{\tilde{q}|B,P_S}(\tilde{q})$ otherwise. Consider what we define as a high tier school S_j: a school where its capacity and the total capacity of all higher ranked school is larger than X_2. Then this school admits all students with perceived quality higher than q_2. Since $f_{\tilde{q}|A,P_S}(\tilde{q}) > f_{\tilde{q}|B,P_S}(\tilde{q})$ on this interval, the proportion τ_j of matched students in group B is smaller than $1 - \pi$ thus group B students are under-represented.

The cases for the remaining intervals follow similarly. $\qquad\square$

Figure 9 illustrates these results. Diversity levels change with tiers[25] due to the joint effects of estimation differences across groups and competition between schools. Differential variances lead to different skill estimate distributions $\tilde{q} \mid g, P_S$ for different groups. In the single school setting, the difference leads to fairness concerns (Proposition 1): more privileged students have perceived skill above the admissions threshold, even conditional on true quality.

In the multiple school setting, the set of students that ultimately attend a school A_j corresponds to a *band* of perceived skill $\tilde{q}$: the students admitted to school A_j but not to any higher ranked school $A_i \succ A_j$. Whether disadvantaged students are under-represented at a school then depends on the mass of each group with perceived skill in that band. Competition across schools leads to different diversity levels, even when all schools choose the same admissions policy.

B.3   One school adopts an intervention

Now we consider a setting where one school deviates from the benchmark model and changes their admission policy. We study the resulting impact on the matching of schools and students. In this setting, one might expect the impact on the system to change depending on whether a low-ranked school or a high-ranked school makes the change. To isolate this effect, we assume that there are two schools A_1 and A_2, where students prefer $A_1 \succ A_2$, and that school A_1 changes their policy to adopt a test-free policy.

Note that when schools observe a different set of features for a student, the schools may have different (though correlated) preferences over the students. In the following proposition, we show that the adoption of test-free policies by the top-rank school leads to the decrease in the average true skill of admitted students of each social group but also an increase in the average true skill of admitted students at the low-rank school.

---

[25]The tiers are mutually exclusive but not necessarily exhaustive. Group B students may be under-represented in a school not included in any of the three tiers.

Proposition 6. *Assume $C_1 + C_2 \leq 0.5$. When A_1 uses a test-free policy and A_2 uses a test-based policy, the academic merit of students of group $g \in \{A, B\}$ decreases at school A_1 and increases at school A_2, compared to when both schools use a test-based policy:*

$$\mathbb{E}[q \mid Y_1 = 1, g, (P_{\text{SUB}}, P_{\text{FULL}})] \leq \mathbb{E}[q \mid Y_1 = 1, g, (P_{\text{FULL}}, P_{\text{FULL}})]$$
$$\mathbb{E}[q \mid Y_2 = 1, g, (P_{\text{SUB}}, P_{\text{FULL}})] \geq \mathbb{E}[q \mid Y_2 = 1, g, (P_{\text{FULL}}, P_{\text{FULL}})].$$

Proof. The result for school A_1 follows directly from part (iii) in Theorem 2.

Let $\boldsymbol{\eta}$ be the set of features $k = 1$ to $K - 1$. For school A_2, from the distribution of $q \mid \tilde{q}, g, P_{a,d}$ for $d = K - 1, K$ and the fact that A_1 has the same capacity C_1 in both scenarios, it follows that, for any constant $c > 0$ and group g,

$$\mathbb{P}[q > c \mid \tilde{q}(\boldsymbol{\theta}, g) < \tilde{q}_{A_1}^*((P_{\text{FULL}}, P_{\text{FULL}}))] \leq \mathbb{P}[q > c \mid \tilde{q}(\boldsymbol{\eta}, g) < \tilde{q}_{A_1}^*((P_{\text{SUB}}, P_{\text{FULL}}))]$$

since the two Normal distributions have the same mean and the variance increases for $K - 1$.

Since school A_2 has the same policy in both the baseline setting and this scenario, the last equation further implies that that

$$\tilde{q}_{A_2}^*((P_{a,\text{SUB}}, P_{\text{FULL}})) \geq \tilde{q}_{A_2}^*((P_{\text{FULL}}, P_{\text{FULL}})).$$

Consequently, A_2 sees an increase in the academic merit of admitted students in g, that is

$$\mathbb{E}[q \mid \tilde{q}_{A_1}^*((P_{\text{SUB}}, P_{\text{FULL}})) > \tilde{q}(\boldsymbol{\theta}, g) \geq \tilde{q}_{A_2}^*((P_{\text{SUB}}, P_{\text{FULL}}))]$$
$$\geq \mathbb{E}[q \mid \tilde{q}_{A_1}^*((P_{\text{SUB}}, P_{\text{FULL}})) > \tilde{q}(\boldsymbol{\theta}, g) \geq \tilde{q}_{A_2}^*((P_{\text{FULL}}, P_{\text{FULL}}))]$$
$$\geq \mathbb{E}[q \mid \tilde{q}_{A_1}^*((P_{\text{FULL}}, P_{\text{FULL}})) > \tilde{q}(\boldsymbol{\theta}, g) \geq \tilde{q}_{A_2}^*((P_{\text{FULL}}, P_{\text{FULL}}))]$$

where the last inequality follows from the academic merit decrease that A_1 experiences for each group g, i.e.,

$$\mathbb{E}[q \mid \tilde{q}(\boldsymbol{\eta}, g) \geq \tilde{q}_{A_1}^*((P_{\text{SUB}}, P_{\text{FULL}}))] \leq \mathbb{E}[q \mid \tilde{q}(\boldsymbol{\theta}, g) \geq \tilde{q}_{A_1}^*((P_{\text{FULL}}, P_{\text{FULL}}))].$$

$\square$

Note that the admission outcomes depend on the vector of policies chosen by both schools, and hence the expectation over true skill is conditioned on this vector.

In this scenario, A_1 drops the test score and is less effective at identifying the top students. This change decreases the average skill level of the students who are admitted to (and thus, the students who attend) A_1. Thus, after A_1 drops the test score, there are high-skilled students that would have been admitted under a test-based policy that are no longer admitted to A_1. These high-skilled students will now attend A_2, if accepted, increasing the expected skill at A_2.

This discussion illustrates how a change made by one school can have downstream effects on aca-

demic merit at other schools. These results are preliminary, but suggestive of possible applications of our model to study the multiple school context.

C   Proofs of Statements

In this appendix, we provide and prove the full statement of each result appearing in the main text.

C.1   Auxiliary lemmas

Let Φ denote the CDF of $\mathcal{N}(0,1)$ and $\mathrm{HR}(x) = \frac{\phi(x)}{1-\Phi(x)}$ the *Hazard Rate* of $X \sim \mathcal{N}(0,1)$.

Lemma C.1. *Let $X \mid M \sim \mathcal{N}(M, \sigma^2)$ and $M \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Then, $X \sim \mathcal{N}(\mu_0, \sigma^2 + \sigma_0^2)$.*

Lemma C.2. *Let $X \mid M \sim \mathcal{N}(M, \sigma^2)$ and $M \sim \mathcal{N}(\mu_0, \sigma_0^2)$. Then,*

$$M \mid X \sim \mathcal{N}\left(\frac{\sigma_0^2}{\sigma^2 + \sigma_0^2} X + \frac{\sigma^2}{\sigma^2 + \sigma_0^2} \mu_0, \frac{1}{\sigma^{-2} + \sigma_0^{-2}} \right).$$

Lemma C.3. *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then, for any $a \in \mathbb{R}$, $\mathbb{E}[X \mid X > a] = \mu + \sigma \frac{\phi(t)}{1-\Phi(t)}$, where $t = \frac{a-\mu}{\sigma}$.*

Lemma C.4. *The hazard rate $\mathrm{HR}(x) = \frac{\phi(x)}{1-\Phi(x)}$, $x \in \mathbb{R}$ has the following properties:*

(i) *Its derivative equals $\frac{\mathrm{dHR}(x)}{\mathrm{d}x} = \mathrm{HR}(x)(\mathrm{HR}(x) - x)$;*

(ii) *It holds that $\mathrm{HR}(x) > x$ for all $x > 0$;*

Lemma C.5. *Let $a > 0$. The function $h(x) = \frac{x}{a}\mathrm{HR}(\frac{a}{x})$ is increasing in $x > 0$.*

Proof. Let $y = a/x$. We study the monotonicity of $\hat{h}(y) = \mathrm{HR}(y)/y$. The derivative of $\hat{h}(y)$ equals

$$\frac{\mathrm{d}\hat{h}(y)}{\mathrm{d}y} = \frac{\frac{\mathrm{dHR}(y)}{\mathrm{d}y} y - \mathrm{HR}(y)}{y^2}.$$

For any $y > 0$, it holds that $\frac{\mathrm{d}\hat{h}(y)}{\mathrm{d}y} < 0$ if and only if $\frac{\mathrm{dHR}(y)}{\mathrm{d}y} y - \mathrm{HR}(y) < 0$. Using part (i) in Lemma C.4, we get that

$$\frac{\mathrm{dHR}(y)}{\mathrm{d}y} y - \mathrm{HR}(y) = \mathrm{HR}(y)\left(\mathrm{HR}(y)\, y - y^2 - 1 \right),$$

which is negative for $y > 0$ if and only if $\mathrm{HR}(y)\, y - y^2 - 1 < 0$ for all $y > 0$.

By Theorem 2.3 in (Baricz, 2008), we know that $\mathrm{HR}(y) < \frac{y}{2} + \frac{\sqrt{y^2+4}}{2}$. Thus, using this inequality, we can bound the quantity $\mathrm{HR}(y)\, y - y^2 - 1$ as follows:

$$\mathrm{HR}(y)\, y - y^2 - 1 < \frac{y^2}{2} + y\frac{\sqrt{y^2+4}}{2} - y^2 - 1 = \frac{y}{2}(-y + \sqrt{y^2+4}) - 1,$$

38

which is negative for any $y \in \mathbb{R}$. Therefore, $\frac{d\hat{h}(y)}{dy} < 0$ for all $y > 0$. Finally, since $\hat{h}(y)$ is decreasing in $y > 0$ and $y = \frac{a}{x}$, $a > 0$, is decreasing in $x > 0$, it follows that $h(x) = \hat{h}\left(\frac{a}{x}\right)$ is increasing in $x > 0$. $\qquad\square$

C.2 Group-aware estimation

Gaussian social learning with feature set $S \subseteq \{1, \ldots, K\}$. Given that $q \sim \mathcal{N}(\mu, \sigma^2)$, $\epsilon_{kg} \sim \mathcal{N}(\mu_{gk}, \sigma_{gk}^2)$ and the noise is drawn independently, each feature $k \in S$ is also normally distributed conditional on q, i.e., $\theta_k \mid q, g \sim \mathcal{N}(q + \mu_{gk}, \sigma_{gk}^2)$. Then, we inductively find that $q \mid \boldsymbol{\theta}, g \sim \mathcal{N}\left(\tilde{q}(\boldsymbol{\theta}, g), \tilde{\sigma}^2(\boldsymbol{\theta}, g)\right)$, where

$$\tilde{q}(\boldsymbol{\theta}, g) = \frac{\mu\sigma^{-2} + \sum_{k \in S}(\theta_k - \mu_{gk})\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{gk}^{-2}}, \quad \tilde{\sigma}^2(\boldsymbol{\theta}, g) = \frac{1}{\sigma^{-2} + \sum_{k \in S}\sigma_{gk}^{-2}}. \tag{7}$$

Perceived skill conditional on true skill. (7) gives us the skill estimate $\tilde{q}$ of a student conditional on features $\boldsymbol{\theta}$. Another useful distribution is the skill estimate conditional on the student's true skill q and group g, i.e., $\tilde{q} \mid q, g, P_S$, which is also Gaussian. Indeed, observe that $\tilde{q}(\boldsymbol{\theta}, g)$ in (7) is a linear combination of independent (conditional on q) Gaussian variables $\theta_k = q + \epsilon_{kg}$, $k \in S$. Thus,

$$\tilde{q} \mid q, g, P_S \sim \mathcal{N}\left(\frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{gk}^{-2}}, \frac{\sum_{k \in S}\sigma_{gk}^{-2}}{\left(\sigma^{-2} + \sum_{k \in S}\sigma_{gk}^{-2}\right)^2}\right). \tag{8}$$

Lemma C.6. *For group-aware estimation policies, the following properties hold:*

(i) $\mathbb{E}[\tilde{q} \mid q, A, P_S] > \mathbb{E}[\tilde{q} \mid q, B, P_S]$ *if and only if* $(q - \mu)\left(\sum_{k \in S}\sigma_{Ak}^{-2} - \sum_{k \in S}\sigma_{Bk}^{-2}\right) > 0$.

(ii) $Var[\tilde{q} \mid q, A, P_S] > Var[\tilde{q} \mid q, B, P_S]$ *if and only if*

$$\left(\sigma^{-4} - \sum_{k \in S}\sigma_{Ak}^{-2}\sum_{k \in S}\sigma_{Bk}^{-2}\right)\left(\sum_{k \in S}\sigma_{Ak}^{-2} - \sum_{k \in S}\sigma_{Bk}^{-2}\right) > 0.$$

Proof. The proof follows immediately from simple algebra thus it is ommitted. $\qquad\square$

Distribution of skill estimates per group. We find the distribution $\tilde{q} \mid g, P_S$, that we denote by $F_{\tilde{q}|g,P_S}$.

Lemma C.7 (Lemma 1). *Consider a school that uses feature set $S \subseteq \{1, \ldots, K\}$ for each applicant. For $g \in \{A, B\}$, the skill level estimates for students in group g are normally distributed:*

$$\tilde{q} \mid g, P_S \sim \mathcal{N}\left(\mu, \sigma^2\left[\frac{\sum_{k \in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{gk}^{-2}}\right]\right).$$

Proof. An application of Lemma C.1 for $X = \tilde{q}$ and $M = \frac{\mu\sigma^{-2} + q\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}$ gives us the result. Analytically, the parameters of this distribution can be computed as follows:

$$\mathbb{E}[\tilde{q} \mid g, P_S] = \mathbb{E}_q[\mathbb{E}[\tilde{q} \mid q, g, P_S]] = \frac{\mu\sigma^{-2} + \mu\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}} = \mu,$$

$$\text{Var}[\tilde{q} \mid g, P_S] = \mathbb{E}[\tilde{q}^2 \mid g, P_S] - \mu^2 = \mathbb{E}_q[\mathbb{E}[\tilde{q}^2 \mid q, g, P_S]] - \mu^2$$

$$= \mathbb{E}_q\left[\text{Var}[\tilde{q} \mid q, g, P_S] + \left(\frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right)^2\right] - \mu^2$$

$$= \mathbb{E}_q[\text{Var}[\tilde{q} \mid q, g, P_S]] + \text{Var}\left[\frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right]$$

$$= \frac{\sum_{k\in S}\sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2})^2} + \sigma^2\left(\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right)^2$$

$$= \sigma^2\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}.$$

$\square$

Corollary 1. *$\text{Var}[\tilde{q} \mid A, P_S] > \text{Var}[\tilde{q} \mid B, P_S]$ if and only if $\sum_{k\in S}\sigma_{Ak}^{-2} > \sum_{k\in S}\sigma_{Bk}^{-2}$.*

Definition 1. *Let X and Y be two random variables with support $\mathbb{R}$ and CDFs F and G, respectively. We say that X second-order stochastically dominates Y, $X \succ_{SSD} Y$, if for every $t \in \mathbb{R}$, $\int_{-\infty}^{t} G(y)dy \geq \int_{-\infty}^{t} F(x)dx$.*

Corollary 2 (Second-order stochastic dominance). *If $\sum_{k\in S}\sigma_{Ak}^{-2} > \sum_{k\in S}\sigma_{Bk}^{-2}$, then $(\tilde{q} \mid B, P_S) \succ_{SSD} (\tilde{q} \mid A, P_S)$ and $\tilde{q} \mid A, P_S$ is a mean-preserving spread of $\tilde{q} \mid B, P_S$.*

Distribution of true skill conditional on skill estimate. To answer questions about the academic merit of the admitted student body, we need to be able to compute the expected value of q conditional on acceptance and the social group g of a student, i.e., $\mathbb{E}[q \mid Y = 1, g, P_S]$. Thus, we first the conditional distribution $q \mid \tilde{q}, g, P_S$ in the following lemma.

Lemma C.8. *Suppose that the school uses policy P_S. Then, the true skill level q of students in group $g \in \{A, B\}$ conditional on the estimated skill level $\tilde{q}$ is normally distributed as follows*

$$q \mid \tilde{q}, g, P_S \sim \mathcal{N}\left(\tilde{q}, \frac{1}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right). \tag{9}$$

Proof. We apply Lemma C.2 by using the transformation $M = \frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_k\sigma_{gk}^{-2}}$ and $X = \tilde{q}$. More specifically, let

$$X \mid M \sim \mathcal{N}\left(M, \frac{\sum_{k\in S}\sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2})^2}\right), \quad M \sim \mathcal{N}\left(\mu, \sigma^2\left(\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right)^2\right).$$

Then, by Lemma C.2, we get that

$$\mathbb{E}[M \mid \tilde{q}, g, P_S] = \frac{\sigma^2 \frac{(\sum_{k \in S} \sigma_{gk}^{-2})^2}{(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2})^2} \tilde{q} + \mu \frac{\sum_{k \in S} \sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2})^2}}{\sigma^2 \frac{(\sum_{k \in S} \sigma_{gk}^{-2})^2}{(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2})^2} + \frac{\sum_{k \in S} \sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2})^2}}$$

$$= \frac{\sum_{k \in S} \sigma_{gk}^{-2} \tilde{q} + \mu \sigma^{-2}}{\sum_{k \in S} \sigma_{gk}^{-2} + \sigma^{-2}}$$

$$\mathrm{Var}[M \mid \tilde{q}, g, P_S] = \left(\left(\frac{\sum_{k \in S} \sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2})^2} \right)^{-1} + \sigma^{-2} \left(\frac{\sum_{k \in S} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2}} \right)^{-2} \right)^{-1}$$

$$= \frac{\left(\sum_{k \in S} \sigma_{gk}^{-2} \right)^2}{\left(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2} \right)^3}.$$

Therefore, $M \mid \tilde{q}, g, P_S \sim \mathcal{N}\left(\frac{\sum_{k \in S} \sigma_{gk}^{-2} \tilde{q} + \mu \sigma^{-2}}{\sum_{k \in S} \sigma_{gk}^{-2} + \sigma^{-2}}, \frac{\left(\sum_{k \in S} \sigma_{gk}^{-2} \right)^2}{\left(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2} \right)^3} \right)$. Finally, using the linear transformation

$$q = \frac{M \left(\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2} \right) - \mu \sigma^{-2}}{\sum_{k \in S} \sigma_{gk}^{-2}},$$

we get that $q \mid \tilde{q}, g, P_S \sim \mathcal{N}\left(\tilde{q}, \frac{1}{\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2}} \right)$. $\qquad\square$

C.3 Baseline policy in the absence of barriers

Let $\tilde{q}_S^*$ denote the optimal decision threshold used by the school under policy P_S. Using the distribution $F_{\tilde{q}|g,P_S}$, it follows that threshold $\tilde{q}_S^*$ is the solution to the equation

$$(1 - \pi) F_{\tilde{q}|A,P_S}(\tilde{q}_S^*) + \pi F_{\tilde{q}|B,P_S}(\tilde{q}_S^*) = 1 - C. \tag{10}$$

By Lemma 1, the *Gaussian mixture* of $F_{\tilde{q}|A,P_S}$, $F_{\tilde{q}|B,P_S}$ with weights $1 - \pi$, π has mean μ and variance

$$(1 - \pi)\sigma^2 \left[\frac{\sum_{k \in S} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Ak}^{-2}} \right] + \pi \sigma^2 \left[\frac{\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}} \right].$$

Recall that for a Gaussian random variable $X \sim N(\mu_0, \sigma_0^2)$, it holds that $\frac{X - \mu_0}{\sigma_0} \sim N(0, 1)$. Thus, (10) can be equivalently written as

$$\Phi \left((\tilde{q}_S^* - \mu) \left((1 - \pi)\sigma^2 \left[\frac{\sum_{k \in S} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Ak}^{-2}} \right] + \pi \sigma^2 \left[\frac{\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}} \right] \right)^{-1/2} \right) = 1 - C. \tag{11}$$

We also introduce some additional definitions. Given any fixed value of $\sum_{k \in S} \sigma_{Bk}^{-2}$, the *informativeness gap* Δ is defined as $\Delta = \sum_{k \in S} \sigma_{Ak}^{-2} - \sum_{k \in S} \sigma_{Bk}^{-2}$. Given all parameters, except $\sum_{k \in S} \sigma_{Ak}^{-2}$

41

fixed, let $F_{\tilde{q}|g,P_S}(q;\Delta)$ denote the CDF $F_{\tilde{q}|g,P_S}$ parameterized by $\Delta \geq 0$ and $\tilde{q}_S^*(\Delta)$ and $\tau(P_S;\Delta)$ denote the corresponding admission threshold and diversity level, respectively, for any $\Delta \geq 0$ under the baseline policy P_S.

We now provide the proof to Proposition 1. Note that the result below considers a general feature set S where the assumption on unequal precisions holds.

Proposition 1 (Metrics with a fixed policy). *Suppose that a selective school uses admissions policy P_S. Group fairness and individual fairness fail except for equal precision. Given unequal precisions:*

(i) Diversity level: *Group B students are under-represented, i.e., $\tau(P_S) < \pi$.*

Furthermore, larger informativeness gap leads to decreased diversity: fix group B precision, $\sum_{k \in S} \sigma_{Bk}^{-2}$; then as group A precision increases, the diversity level $\tau(P_S)$ decreases.

(ii) Individual fairness: *High-skilled group B students are hard to target, i.e., $I(q;P_S) > 0$, if and only if*

$$q > \tilde{q}_S^* + \frac{\sigma^{-2}(\tilde{q}_S^* - \mu)}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}.$$

Increasing the informativeness gap increases the individual fairness gap for high-skilled students: fix group B precision, $\sum_{k \in S} \sigma_{Bk}^{-2}$; then as group A precision increases, $I(q;P_S)$ increases for $q > \mu + \sigma\Phi^{-1}(1 - C)$.

(iii) Academic merit: *The policy achieves worse academic merit for admitted students from group B.*

Proof of part (i). We break the proof into two steps.

Step 1: We show that group fairness fails except for equal precision. Given unequal precisions, we further show that $\tau(P_S) < \pi$. If $\sum_{k \in S} \sigma_{Ak}^{-2} = \sum_{k \in S} \sigma_{Bk}^{-2}$, then the two distributions $F_{\tilde{q}|A,P_S}$, $F_{\tilde{q}|B,P_S}$ are identical so it trivially holds that $F_{\tilde{q}|A,P_S}(\tilde{q}_S^*) = F_{\tilde{q}|A,P_S}(\tilde{q}_S^*) = 1 - C$. Consequently, group fairness is achieved.

Next, assume that $\sum_{k \in S} \sigma_{Ak}^{-2} > \sum_{k \in S} \sigma_{Bk}^{-2}$. Then, by Lemma 1 and Corollary 2, $(\tilde{q} \mid B, P_S) \succ_{SSD} (\tilde{q} \mid A, P_S)$ and $\tilde{q} \mid A, P_S$ is a mean-preserving spread of $\tilde{q} \mid B, P_S$. Thus, the CDFs $F_{\tilde{q}|A,P_S}$ and $F_{\tilde{q}|B,P_S}$ cross once at $\tilde{q} = \mu$. Furthermore, $F_{\tilde{q}|A,P_S}(\tilde{q}) < F_{\tilde{q}|B,P_S}(\tilde{q})$, for $\tilde{q} > \mu$ and $F_{\tilde{q}|A,P_S}(\tilde{q}) > F_{\tilde{q}|B,P_S}(\tilde{q})$, for $\tilde{q} < \mu$.

Since $C < 0.5 = F_{\tilde{q}|A,P_S}(\mu) = F_{\tilde{q}|B,P_S}(\mu)$, then $\tilde{q}_S^* > \mu$. Therefore, $F_{\tilde{q}|A,P_S}(\tilde{q}_S^*) < F_{\tilde{q}|B,P_S}(\tilde{q}_S^*)$, which due to (10) implies that $1 - F_{\tilde{q}|B,P_S}(\tilde{q}_S^*) < C$ thus

$$\tau(P_S) = \frac{\pi(1 - F_{\tilde{q}|B,P_S}(\tilde{q}_S^*))}{C} < \pi.$$

Step 2: We show that the marginal effect of Δ on $\tau(P_S)$ is negative. Consider $0 \leq \Delta < \Delta'$. Since $F_{\tilde{q}|B,P_S}(q;\Delta)$ depends only on $\sum_{k \in S} \sigma_{Bk}^{-2}$, it remains unchanged under both Δ, Δ'.

Recall that the admission threshold is the solution to (11). Solving for $\tilde{q}_S^*(\Delta)$ gives us

$$\tilde{q}_S^*(\Delta) = \mu + \Phi^{-1}(1-C) \cdot \left((1-\pi)\sigma^2 \left[\frac{\sum_{k\in S}\sigma_{Bk}^{-2} + \Delta}{\sigma^{-2} + \sum_{k\in S}\sigma_{Bk}^{-2} + \Delta} \right] + \pi\sigma^2 \left[\frac{\sum_{k\in S}\sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{Bk}^{-2}} \right] \right)^{1/2},$$
(12)

which is an increasing function of Δ. Thus, $\tilde{q}_S^*(\Delta') > \tilde{q}_S^*(\Delta)$.

Therefore, given that the capacity remains constant at C, the diversity level decreases as Δ increases since

$$\tau(P_S; \Delta') = \frac{\pi(1 - F_{\tilde{q}|B,P_S}(\tilde{q}_S^*(\Delta'); \Delta'))}{C} = \frac{\pi(1 - F_{\tilde{q}|B,P_S}(\tilde{q}_S^*(\Delta'); \Delta))}{C} < \frac{\pi(1 - F_{\tilde{q}|B,P_S}(\tilde{q}_S^*(\Delta); \Delta))}{C} = \tau(P_S; \Delta).$$

$\square$

Proof of part (ii). We prove each claim in different steps.

Step 1: We show that $I(q; P_S) > 0$ if and only if

$$q > \tilde{q}_S^* + \frac{\sigma^{-2}(\tilde{q}_S^* - \mu)}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}}.$$

Recall that for a Gaussian variable $X \sim N(\mu_0, \sigma_0^2)$, it holds that $\frac{X-\mu_0}{\sigma_0} \sim N(0,1)$. Thus, given policy P_S, the probability of admission for a student in group g equals

$$\mathbb{P}[Y = 1 \mid q, g, P_S] = 1 - F_{\tilde{q}|q,g,P_S}(\tilde{q}_S^*) = 1 - \Phi\left(\frac{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{gk}^{-2}}} \left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}} \right) \right),$$
(13)

where

$$\mathbb{E}[\tilde{q} \mid q, g, P_S] = \frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}, \mathrm{Var}[\tilde{q} \mid q, g, P_S] = \frac{\sum_{k\in S}\sigma_{gk}^{-2}}{\left(\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}\right)^2}.$$

Consequently, due to the monotonicity of Φ, it holds that $I(q; P_S) > 0$ if and only if

$$\frac{\sigma^{-2} + \sum_{k\in S}\sigma_{Ak}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}} \left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{Ak}^{-2}} \right) < \frac{\sigma^{-2} + \sum_{k\in S}\sigma_{Bk}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}} \left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k\in S}\sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{Bk}^{-2}} \right)$$

$$\iff \frac{\tilde{q}_S^*\sigma^{-2} + \tilde{q}_S^*\sum_{k\in S}\sigma_{Ak}^{-2} - \mu\sigma^{-2} - q\sum_{k\in S}\sigma_{Ak}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}} < \frac{\tilde{q}_S^*\sigma^{-2} + \tilde{q}_S^*\sum_{k\in S}\sigma_{Bk}^{-2} - \mu\sigma^{-2} - q\sum_{k\in S}\sigma_{Bk}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}}$$

$$\iff \left(\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}} - \sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} \right) \left(\sigma^{-2}(\tilde{q}_S^* - \mu) + (q - \tilde{q}_S^*)\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} \right) < 0.$$
(14)

Due to our assumption on unequal precisions, the last inequality further translates to

$$(\tilde{q}_S^* - q)\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} < \sigma^{-2}(\tilde{q}_S^* - \mu),$$

where the RHS is always positive due to school selectivity which implies that $\tilde{q}_S^* > \mu$. Thus, we conclude that $I(q; P_S) > 0$ if and only if

$$q > \tilde{q}_S^* + \frac{\sigma^{-2}(\tilde{q}_S^* - \mu)}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}.$$

Step 2: We show that individual fairness fails except for equal precisions. As an immediate corollary of the previous analysis in Step 1, observe that individual fairness fails unless the LHS in (14) equals 0 for all q; equivalently, individual fairness fails except for equal precision, i.e., $\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}} - \sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} = 0$.

Step 3: Finally, we show that for $q > \mu + \sigma\Phi^{-1}(1 - C)$, $I(q; P_S)$ increases as the informativeness gap increases. We begin with group B. By (8), it follows that

$$\mathbb{P}[Y = 1 \mid q, B, P_S, \Delta] = 1 - F_{\tilde{q}|B,P_S}(\tilde{q}_S^*(\Delta); \Delta) = 1 - \Phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}\left(\tilde{q}_S^*(\Delta) - \frac{\mu\sigma^{-2} + q\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}}\right)\right).$$

By (12), it further follows that $\tilde{q}_S^*(\Delta)$ is increasing in Δ. Consequently, the above probability is decreasing in Δ since Φ is an increasing function and all terms except for $\tilde{q}_S^*(\Delta)$ do not depend on Δ. Therefore, we conclude that the admission probability of group B students decreases for any q as Δ increases.

Next, for group A, note that students with $q > \mu + \sigma\Phi^{-1}(1 - C)$ are exactly those students in group A who – given perfectly observable skills q – would be admitted to the class; due to imperfect information, a group A student of true skill $q > \mu + \sigma\Phi^{-1}(1 - C)$ has a non-zero probability to get rejected. Next, observe that as Δ increases, the total precision $\sum_{k \in S} \sigma_{Ak}^{-2}$ of group A must increase. Consequently, the variance $\text{Var}[\tilde{q} \mid q, A, P_S]$ decreases thus the estimates $\tilde{q} \mid q, A, P_S$ of all group A students (including those with true skill $q > \mu + \sigma\Phi^{-1}(1 - C)$) become more precise. Combining this observation with the facts that the capacity C remains constant and the admission probability of group B students decreases, it follows that the probability that the top-skilled group A students with $q > \mu + \sigma\Phi^{-1}(1 - C)$ are rejected (either in favor of lower-skilled students in A or students in B) decreases as Δ increases. Equivalently, their admission probability $\mathbb{P}[Y = 1 \mid q, A, P_S, \Delta]$ increases as Δ grows.

Putting everything together, we conclude that, given $q > \mu + \sigma\Phi^{-1}(1 - C)$, the individual fairness gap $I(q; P_S)$ increases as the informativeness gap Δ increases. $\qquad\square$

Proof of part (iii). We break the proof into the following steps.

Step 1: *We compute the expected value* $\mathbb{E}[\tilde{q} \mid Y = 1, g, P_S]$ *and show that* $\mathbb{E}[\tilde{q} \mid Y = 1, A, P_S] \geq$ $\mathbb{E}[\tilde{q} \mid Y = 1, B, P_S]$. *Applying Lemma C.3, we get that*

$$
\begin{aligned}
\mathbb{E}[\tilde{q} \mid Y = 1, g, P_S] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_S^*, g, P_S] &= \mathbb{E}[\tilde{q} \mid g] + \sqrt{\mathrm{Var}[\tilde{q} \mid g, P_S]} \frac{\phi(t_g)}{1 - \Phi(t_g)} \\
&= \mu + \sigma \sqrt{\frac{\sum_{k \in S} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2}}} \cdot \frac{\phi(t_g)}{1 - \Phi(t_g)},
\end{aligned}
\tag{15}
$$

where $t_g = \frac{\tilde{q}_S^* - \mathbb{E}[\tilde{q}|g, P_S]}{\sqrt{\mathrm{Var}[\tilde{q}|g, P_S]}}$. Due to school selectivity, we have $\tilde{q}_S^* > \mu$. By Lemma C.5, the function

$$
h(x) = x \frac{\phi\left(\frac{\tilde{q}_S^* - \mu}{x}\right)}{1 - \Phi\left(\frac{\tilde{q}_S^* - \mu}{x}\right)} = x \mathrm{HR}\left(\frac{\tilde{q}_S^* - \mu}{x}\right)
$$

is increasing in $x > 0$ for $\tilde{q}_S^* > \mu$. Thus, by Corollary 1, we get that that

$$
\mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_S^*, A, P_S] \geq \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_S^*, B, P_S].
$$

Step 2: *We compute the expected value* $\mathbb{E}[q \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S]$. *Specifically,*

$$
\mathbb{E}[q \mid Y = 1, g, P_S] = \mathbb{E}[q \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S] = \mathbb{E}_{\tilde{q}}[\mathbb{E}_q[q \mid \tilde{q}, g, P_S] \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S],
\tag{16}
$$

where the last equality follows from Lemma C.8.

Step 3: *We show that* $\mathbb{E}[q \mid Y = 1, A, P_S] > \mathbb{E}[q \mid Y = 1, B, P_S]$. Given our assumptions on unequal precisions and school selectivity, the proof follows immediately from Steps 1 and 2. I.e., if $\sum_{k \in S} \sigma_{Ak}^{-2} > \sum_{k \in S} \sigma_{Bk}^{-2}$ and $C < 0.5$, then $\mathbb{E}[q \mid Y = 1, A, P_S] > \mathbb{E}[q \mid Y = 1, B, P_S]$. $\qquad\square$

Explaining why the individual fairness gap decreases for high-skilled students. We include an observation about the individual fairness gap. Although the individual fairness gap is positive for sufficiently high-skilled students, the magnitude of this gap varies. For students at the end of the right tail of the true skill distribution, the individual fairness gap starts to decrease. This property can be graphically observed in Figure 4b.

Lemma 2. *Consider policy P_S, and assume unequal precision. The individual fairness gap $I(q; P_S)$ is decreasing in q for $q > q_e$, where*

$$
q_e \triangleq \tilde{q}_S^* + \sqrt{\frac{\sigma^{-4}(\mu - \tilde{q}_S^*)^2}{\sum_{k \in S} \sigma_{Ak}^{-2} \sum_{k \in S} \sigma_{Bk}^{-2}} + \frac{\ln\left(\sum_{k \in S} \sigma_{Ak}^{-2}\right) - \ln\left(\sum_{k \in S} \sigma_{Bk}^{-2}\right)}{\sum_{k \in S} \sigma_{Ak}^{-2} - \sum_{k \in S} \sigma_{Bk}^{-2}}}.
$$

Furthermore, $\lim_{q \to \infty} I(q; P_S) = 0$.

Proof. By (8), the individual fairness gap equals

$$I(q; P_S) = \left(1 - \Phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}\left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{Ak}^{-2}}\right)\right)\right)$$
$$-\left(1 - \Phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}\left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{Bk}^{-2}}\right)\right)\right).$$

Taking the derivative of $I(q; P_S)$ with respect to q, we find that

$$\frac{\mathrm{d}I(q; P_S)}{\mathrm{d}q} = \phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}\left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{Ak}^{-2}}\right)\right)\sqrt{\sum_{k \in S}\sigma_{Ak}^{-2}}$$
$$- \phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}\left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{Bk}^{-2}}\right)\right)\sqrt{\sum_{k \in S}\sigma_{Bk}^{-2}}.$$

Thus, to prove that $\frac{\mathrm{d}I(q; P_S)}{\mathrm{d}q} < 0$, it suffices to show that

$$\ln\left(\phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}\left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{Ak}^{-2}}\right)\right)\sqrt{\sum_{k \in S}\sigma_{Ak}^{-2}}\right)$$
$$< \ln\left(\phi\left(\frac{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}\left(\tilde{q}_S^* - \frac{\mu\sigma^{-2} + q\sum_{k \in S}\sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S}\sigma_{Bk}^{-2}}\right)\right)\sqrt{\sum_{k \in S}\sigma_{Bk}^{-2}}\right).$$

The above condition is equivalent to

$$-\frac{\left((\tilde{q}_S^* - \mu)\sigma^{-2} + (\tilde{q}_S^* - q)\sum_{k \in S}\sigma_{Ak}^{-2}\right)^2}{\sum_{k \in S}\sigma_{Ak}^{-2}} + \ln\left(\sum_{k \in S}\sigma_{Ak}^{-2}\right) < -\frac{\left((\tilde{q}_S^* - \mu)\sigma^{-2} + (\tilde{q}_S^* - q)\sum_{k \in S}\sigma_{Bk}^{-2}\right)^2}{\sum_{k \in S}\sigma_{Bk}^{-2}} + \ln\left(\sum_{k \in S}\sigma_{Bk}^{-2}\right)$$
$$\iff \left(\sum_{k \in S}\sigma_{Ak}^{-2} - \sum_{k \in S}\sigma_{Bk}^{-2}\right)\left(\sigma^{-4}(\mu - \tilde{q}_S^*)^2 - \sum_{k \in S}\sigma_{Ak}^{-2}\sum_{k \in S}\sigma_{Bk}^{-2}(q - \tilde{q}_S^*)^2\right) + \sum_{k \in S}\sigma_{Ak}^{-2}\sum_{k \in S}\sigma_{Bk}^{-2}\ln\left(\frac{\sum_{k \in S}\sigma_{Ak}^{-2}}{\sum_{k \in S}\sigma_{Bk}^{-2}}\right) < 0.$$

Given our assumption on unequal precision, i.e., $\sum_{k \in S}\sigma_{Bk}^{-2} < \sum_{k \in S}\sigma_{Ak}^{-2}$, we further get that this condition is satisfied for

$$q > q_e \triangleq \tilde{q}_S^* + \sqrt{\frac{\sigma^{-4}(\mu - \tilde{q}_S^*)^2}{\sum_{k \in S}\sigma_{Ak}^{-2}\sum_{k \in S}\sigma_{Bk}^{-2}} + \frac{\ln\left(\sum_{k \in S}\sigma_{Ak}^{-2}\right) - \ln\left(\sum_{k \in S}\sigma_{Bk}^{-2}\right)}{\sum_{k \in S}\sigma_{Ak}^{-2} - \sum_{k \in S}\sigma_{Bk}^{-2}}}.$$

Therefore, the individual fairness gap $I(q; P_S)$ is decreasing in q for $q > q_e$ as desired.

Furthermore, by the definition of $I(q; P_S)$ and the fact that $\lim_{q' \to \infty} \Phi(q') = 1$, we immediately get that $\lim_{q \to \infty} I(q; P_S) = 0$. $\qquad\square$

C.4 Dropping a feature in the absence of barriers

We are interested in comparing group-aware policies P_{FULL} and P_{SUB}. By our previous result in Lemma 1, we get that

$$\tilde{q} \mid g, P_{\text{SUB}} \sim \mathcal{N}\left(\mu, \sigma^2 \frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}\right), \quad \tilde{q} \mid g, P_{\text{FULL}} \sim \mathcal{N}\left(\mu, \sigma^2 \frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}\right).$$

Lemma C.9. *The variance of $\tilde{q} \mid g, P_{\text{SUB}}$ is lower than that of $\tilde{q} \mid g, P_{\text{FULL}}$ but their means are both equal to μ.*

Proof. The proof follows trivially from the fact that the function $h(x) = \frac{x}{\sigma^{-2} + x}$ is increasing in $x > 0$ and

$$\sum_{k \in \text{FULL}} \sigma_{gk}^{-2} = \sum_{k=1}^{K} \sigma_{gk}^{-2} > \sum_{k=1}^{K-1} \sigma_{gk}^{-2} = \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}$$

for any g. $\qquad\square$

Let $\tilde{q}_{\text{SUB}}^*$ be the decision threshold of a school considering only features $k = 1$ to $K - 1$. By (10), $\tilde{q}_{\text{SUB}}^*$ is the solution to the following equation

$$(1 - \pi) F_{\tilde{q} \mid A, P_{\text{SUB}}}(\tilde{q}_{\text{SUB}}^*) + \pi F_{\tilde{q} \mid A, P_{\text{SUB}}}(\tilde{q}_{\text{SUB}}^*) = 1 - C,$$

whereas $\tilde{q}_{\text{FULL}}^*$ is the solution to

$$(1 - \pi) F_{\tilde{q} \mid A, P_{\text{FULL}}}(\tilde{q}_{\text{FULL}}^*) + \pi F_{\tilde{q} \mid A, P_{\text{FULL}}}(\tilde{q}_{\text{FULL}}^*) = 1 - C.$$

Lemma C.10. *The admission threshold decreases after dropping feature $k = K$, i.e., $\tilde{q}_{\text{SUB}}^* < \tilde{q}_{\text{FULL}}^*$.*

Proof. The proof follows from the definitions of $\tilde{q}_{\text{SUB}}^*, \tilde{q}_{\text{FULL}}^*$, and Lemma C.9. $\qquad\square$

Theorem 2 (Dropping tests without barriers). *Consider policies P_{FULL} and P_{SUB}, and assume unequal precisions under P_{FULL}.*

(i) Diversity level: *Diversity level improves after dropping feature K, $\tau(P_{\text{SUB}}) > \tau(P_{\text{FULL}})$, if and only if*

$$\frac{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}\left(\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}\right)}{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}\left(\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}\right)} < \frac{\sigma_{AK}^{-2}}{\sigma_{BK}^{-2}}. \tag{3}$$

(ii) Individual fairness: *For each group g, there exist thresholds q_g such that the admission probability for students of skill q in group g decreases under P_{SUB} if and only if $q > q_g$. Further, there exists a threshold $\hat{q} \geq \max\{q_A, q_B\}$ such that the individual fairness gap increases for all $q > \hat{q}$, but may decrease otherwise.*

(iii) Academic merit: *Academic merit decreases for both groups $g \in \{A, B\}$, i.e.,*

$$\mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}] > \mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}].$$

Proof of part (i). Diversity improves if and only if

$$\tau(P_{\text{SUB}}) = 1 - F_{\tilde{q}|A,P_{\text{SUB}}}(\tilde{q}^*_{\text{SUB}}) > 1 - F_{\tilde{q}|A,P_{\text{FULL}}}(\tilde{q}_a) = \tau(P_{\text{FULL}}).$$

By the definition of diversity level and Lemma 1, this is equivalent to the following condition

$$1 - \Phi\left(\frac{\tilde{q}^*_{\text{SUB}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}+\sigma^{-2}}}}\right) > 1 - \Phi\left(\frac{\tilde{q}^*_{\text{FULL}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}+\sigma^{-2}}}}\right).$$

Replacing $\tilde{q}^*_{\text{FULL}}, \tilde{q}^*_{\text{SUB}}$ with their definitions as in (11), the above inequality becomes

$$\Phi\left(\Phi^{-1}(1-C)\sqrt{(1-\pi)\frac{\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}} + \pi}\right) < \Phi\left(\Phi^{-1}(1-C)\sqrt{(1-\pi)\frac{\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}} + \pi}\right),$$

which – due to the monotonicity of Φ – holds if and only if

$$\frac{\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}} < \frac{\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}}.$$

Using the substitution $\sum_{k\in\text{FULL}}\sigma_{gk}^{-2} = \sum_{k\in\text{SUB}}\sigma_{gk}^{-2} + \sigma_{gK}$, the last relation equivalently simplifies to (3). $\qquad\square$

Proof of part (ii). We prove each claim at a separate step.

Step 1: We show that, for group B, $\mathbb{P}(Y = 1 \mid q, B, P_{\text{FULL}}) < \mathbb{P}(Y = 1 \mid q, B, P_{\text{SUB}})$ if and only if

$$q < q_B \triangleq \mu + \frac{\sigma\Phi^{-1}(1-C)}{\sqrt{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}} - \sqrt{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\left(\sqrt{\sigma^{-2} + \sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}\sqrt{(1-\pi)\frac{\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}} + \pi}\right.$$

$$\left. - \sqrt{\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}\sqrt{(1-\pi)\frac{\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}} + \pi}\right),$$

Similarly, for group A, it holds that $\mathbb{P}(Y = 1 \mid q, A, P_{\text{FULL}}) > \mathbb{P}(Y = 1 \mid q, A, P_{\text{SUB}})$ if and only if

$$q < q_A \triangleq \mu + \frac{\sigma \Phi^{-1}(1-C)}{\sqrt{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}} - \sqrt{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}} \left(\sqrt{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}} \sqrt{(1-\pi) + \pi \frac{\frac{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}}} \right.$$

$$\left. - \sqrt{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}} \sqrt{(1-\pi) + \pi \frac{\frac{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}}{\frac{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}}} \right),$$

Assume $g = B$; the proof for group A is analogous. Replacing $\tilde{q}_S^*$ from (11) in (13), we find that for policy P_S, the admissions probability (conditional on true skill q and group g) equals

$$\mathbb{P}(Y = 1 \mid q, B, P_S) = 1 - \Phi \left((\mu - q)\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}} + \sigma \Phi^{-1}(1-C)\sqrt{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}} \sqrt{\frac{(1-\pi)\frac{\sum_{k \in S} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{Bk}^{-2}} + \pi \frac{\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{Bk}^{-2}}}{\frac{\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{Bk}^{-2}}}} \right).$$

Thus, the admission probability increases after dropping test scores, if and only if

$$(\mu - q) \left(\sqrt{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}} - \sqrt{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}} \right) + \sigma \Phi^{-1}(1-C)\sqrt{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}} \sqrt{(1-\pi)\frac{\frac{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}} + \pi}$$

$$- \sigma \Phi^{-1}(1-C)\sqrt{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}} \sqrt{(1-\pi)\frac{\frac{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}} + \pi} > 0. \tag{17}$$

This is equivalent to $q < q_B$, i.e.,

$$q < \mu + \frac{\sigma \Phi^{-1}(1-C)}{\sqrt{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}} - \sqrt{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}} \left(\sqrt{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}} \sqrt{(1-\pi)\frac{\frac{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}} + \pi} \right.$$

$$\left. - \sqrt{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}} \sqrt{(1-\pi)\frac{\frac{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}}{\frac{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}} + \pi} \right),$$

Step 2: We show that there exists a threshold $\hat{q} \geq \max\{q_A, q_B\}$ such that the individual fairness gap increases for all $q > \hat{q}$. Otherwise, it may decrease. Let

$$\underline{q} \triangleq \arg\min_{q \in \mathbb{R}} \left\{ (\mu - q)\sqrt{\sum_{k \in S} \sigma_{gk}^{-2}} + \sigma \Phi^{-1}(1-C)\sqrt{\sigma^{-2} + \sum_{k \in S} \sigma_{gk}^{-2}} \sqrt{\frac{(1-\pi)\frac{\sum_{k \in S} \sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{Ak}^{-2}} + \pi \frac{\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{Bk}^{-2}}}{\frac{\sum_{k \in S} \sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k \in S} \sigma_{gk}^{-2}}}} \leq 0, \forall g, S \right\}.$$

Next, consider only $q > \max\{\underline{q}, q_A, q_B\}$. Since Φ is monotone and convex in $(-\infty, 0]$ and $q > \hat{q}$, and by Step 1 for any group g, it also holds that $\mathbb{P}(Y = 1 \mid q, g, P_{\text{FULL}}) > \mathbb{P}(Y = 1 \mid q, g, P_{\text{SUB}})$ for

all $q > q_g$, a sufficient condition for $I(q; P_{\text{FULL}}) > I(q; P_{\text{SUB}})$, to hold is

$$\frac{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}} \left(\tilde{q}_{\text{SUB}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}} \right) - \frac{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}} \left(\tilde{q}_{\text{FULL}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}} \right)$$

$$< \frac{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}} \left(\tilde{q}_{\text{SUB}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}} \right) - \frac{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}} \left(\tilde{q}_{\text{FULL}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}} \right).$$

Let

$$\underline{q} \triangleq \arg\min_{q \in \mathbb{R}} \left\{ \frac{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}} \left(\tilde{q}_{\text{SUB}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Ak}^{-2}} \right) - \frac{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}} \left(\tilde{q}_{\text{FULL}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}} \right) \right.$$

$$\left. < \frac{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}} \left(\tilde{q}_{\text{SUB}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{Bk}^{-2}} \right) - \frac{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}} \left(\tilde{q}_{\text{FULL}}^* - \frac{\mu \sigma^{-2} + q \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}} \right) \right\}.$$

Define $\hat{q} \triangleq \max\{\underline{q}, \underline{q}, q_A, q_B\}$. Then, by the previous conditions, we have $I(q; P_{\text{FULL}}) > I(q; P_{\text{SUB}})$ for all $q > \hat{q}$, thus the individual fairness gap decreases. Furthermore, $\hat{q} \geq \max\{q_A, q_B\}$ as required.

Finally, if $q_A < q_B$, then for all $q_A < q < q_B$, $\mathbb{P}(Y = 1 \mid q, A, P_{\text{FULL}}) > \mathbb{P}(Y = 1 \mid A, g, P_{\text{SUB}})$ but $\mathbb{P}(Y = 1 \mid q, B, P_{\text{FULL}}) < \mathbb{P}(Y = 1 \mid B, g, P_{\text{SUB}})$ (by Step 1). Thus, $I(q; P_{\text{FULL}}) > I(q; P_{\text{SUB}})$. $\qquad \square$

Proof of part (iii). Since $\text{Var}[\tilde{q} \mid g, P_{\text{SUB}}] < \text{Var}[\tilde{q} \mid g, P_{\text{FULL}}]$ and, by Corollary C.10, $\tilde{q}_{\text{SUB}}^* < \tilde{q}_{\text{FULL}}^*$, the expected estimated skill of each admitted group decreases, that is

$$\mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{\text{SUB}}^*, g, P_{\text{SUB}}] < \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{\text{FULL}}^*, g, P_{\text{FULL}}].$$

(16) further implies that $\mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}] < \mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}]$. $\qquad \square$

C.5 Admissions with barriers to testing

In a setting with barriers to testing and policy P_{FULL}, let $\tilde{w}_{\text{FULL}}^*$ the decision threshold of the school with policy P_{FULL}. Then, observe that $\tilde{w}_{\text{FULL}}^* < \tilde{q}_{\text{FULL}}^*$, where

$$(1 - \pi)\gamma_A(1 - F_{\tilde{q}|A, P_{\text{FULL}}}(\tilde{w}_{\text{FULL}}^*)) + \pi\gamma_B(1 - F_{\tilde{q}|B, P_{\text{FULL}}}(\tilde{w}_{\text{FULL}}^*)) = C. \tag{18}$$

We now study the trade-off between barriers and informativeness. For brevity, we use $\pi_A = 1 - \pi$, $\pi_B = \pi$.

Theorem 3 (Theorem 1). *Consider policies P_{FULL} and P_{SUB} and assume unequal precisions under P_{FULL}.*

(i) *For each group g there exists a constant $\Delta_g(\xi_g, \rho_{\text{SUB}}^g)$ such that the academic merit of group g increases if and only if*

$$\beta_g(\gamma_A, \gamma_B, \rho_{\text{FULL}}^g) \leq \Delta_g(\xi_g, \rho_{\text{SUB}}^g), \tag{19}$$

where

$$\rho_S^A = \frac{1}{\rho_S^B} \triangleq \frac{\sum_{k \in S} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Bk}^{-2}} \left(\frac{\sum_{k \in S} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in S} \sigma_{Ak}^{-2}} \right)^{-1},$$

$$\xi_g = \frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}} \left(\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}} \right)^{-1},$$

$$\beta_g(\gamma_g, \gamma_{g'}, \rho_S^g) \triangleq \Phi^{-1} \left(1 - \frac{C}{\pi_g \gamma_g + \pi_{g'} \gamma_{g'}} \right) \sqrt{\frac{\frac{\pi_{g'} \gamma_{g'}}{\pi_g \gamma_g} \rho_S^g + 1}{1 + \frac{\pi_{g'} \gamma_{g'}}{\pi_g \gamma_g}}},$$

$$\Delta_g(\xi_g, \rho_{\text{SUB}}^g) = HR^{-1} \left(\xi_g HR \left(\Phi^{-1}(1 - C) \sqrt{\pi_g + \pi_{g'} \rho_{\text{SUB}}^g} \right) \right).$$

As barriers to group g increase (γ_g decreases), then $\beta_g(\gamma_A, \gamma_B, \rho_{\text{SUB}}^g)$ decreases. Thus, given any group g and $\gamma_{g'} \in (0,1]$, $g' \neq g$, there exists threshold $\bar{\bar{\gamma}}_g \in (0,1]$, such that academic merit of group g improves by dropping feature K if and only if $\gamma_g < \bar{\bar{\gamma}}_g$.

(ii) *Diversity strictly improves after dropping test scores if and only if $\eta(1, 1, \rho_{\text{SUB}}^B) > \eta(\gamma_A, \gamma_B, \rho_{\text{FULL}}^B)$, where*

$$\eta(\gamma_A, \gamma_B, \rho_S^B) \triangleq \frac{(1-\pi)\gamma_B}{C} \left(1 - \Phi \left(\Phi^{-1} \left(1 - \frac{C}{(1-\pi)\gamma_A + \pi\gamma_B} \right) \sqrt{\frac{(1-\pi)\gamma_A \rho_S^B + \pi\gamma_B}{(1-\pi)\gamma_A + \pi\gamma_B}} \right) \right).$$

Given any $\gamma_A \in (0,1]$, there exists a threshold $\bar{\gamma} \in (0,1]$, such that diversity strictly improves after dropping test scorers if and only if $\gamma_B < \bar{\gamma}$.

Proof of part (i). We break the proof into the following parts.

Step 1: We show that the academic merit of group g increases if and only if (19) *holds.* We adopt an argument similar to the proof of Proposition 3. We prove the statement for $g = A$. The argument for group B is similar.

First, similarly to (11), we derive that

$$\tilde{w}_{\text{FULL}}^* = \mu + \Phi^{-1} \left(1 - \frac{C}{(1-\pi)\gamma_A + \pi\gamma_B} \right) \sigma \sqrt{\frac{(1-\pi)\gamma_A \frac{\sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Ak}^{-2}} + \pi\gamma_B \frac{\sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{Bk}^{-2}}}{(1-\pi)\gamma_A + \pi\gamma_B}}.$$

$$(20)$$

Second, requiring that $\mathbb{E}[q \mid Y = 1, A, P_{\text{FULL}}] \leq \mathbb{E}[q \mid Y = 1, A, P_{\text{SUB}}]$ and adapting Lemma C.3

to our setting with barriers gives us

$$
\sqrt{\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}}\,\text{HR}\left(\frac{\Phi^{-1}\left(1-\frac{C}{(1-\pi)\gamma_A+\pi\gamma_B}\right)\sqrt{\frac{\frac{(1-\pi)\gamma_A}{\pi\gamma_B}\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}+\frac{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}}{1+\frac{(1-\pi)\gamma_A}{\pi\gamma_B}}}}{\sqrt{\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}}}\right)
$$

$$
\leq\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}}\,\text{HR}\left(\frac{\Phi^{-1}(1-C)\sqrt{(1-\pi)\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}+\pi\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}}{\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}}}\right).
$$

Replacing with the definitions of $\Delta_A, \rho_{\text{FULL}}^A$, we finally obtain that

$$
\Phi^{-1}\left(1-\frac{C}{(1-\pi)\gamma_A+\pi\gamma_B}\right)\sqrt{\frac{\frac{(1-\pi)\gamma_A}{\pi\gamma_B}+\rho_{\text{FULL}}^A}{1+\frac{(1-\pi)\gamma_A}{\pi\gamma_B}}}\leq\Delta_A(\xi_A,\rho_{\text{SUB}}^A).
$$

Equivalently, using the definition of β_g, we finally get that academic merit in group A improves after dropping feature K if and only if $\beta_A(\gamma_A,\gamma_B,\rho_{\text{FULL}}^A)\leq\Delta_A(\xi_A,\rho_{\text{SUB}}^A)$.

Step 2: We show that, for each group $g\in\{A,B\}$, $\beta_g(\gamma_g,\gamma_{g'},\rho_{\text{FULL}}^g)$ is increasing in γ_g. Given some group g, fix all parameters except γ_g. Then, the function $\Phi^{-1}\left(1-\frac{C}{(1-\pi)\gamma_A+\pi\gamma_B}\right)$ is increasing in γ_g since Φ^{-1} is increasing in its argument and $1-\frac{C}{(1-\pi)\gamma_A+\pi\gamma_B}$ is an increasing function of both γ_A,γ_B.

Now consider the expression in the second term of β:

$$
\sqrt{\frac{\frac{\pi_{g'}\gamma_{g'}}{\pi_g\gamma_g}\rho_S^g+1}{1+\frac{\pi_{g'}\gamma_{g'}}{\pi_g\gamma_g}}}. \tag{21}
$$

We show that this function is increasing in γ_g, for both $g=A$ and $g=B$. More specifically, for group $g=A$, the derivative of (21) with respect to γ_A equals

$$
\frac{\partial}{\partial\gamma_A}\left(\sqrt{\frac{\frac{(1-\pi)\gamma_A}{\pi\gamma_B}+\rho_{\text{FULL}}^A}{1+\frac{(1-\pi)\gamma_A}{\pi\gamma_B}}}\right)=\frac{(1-\pi)\pi\gamma_B(1-\rho_{\text{FULL}}^A)}{2((1-\pi)\gamma_A+\pi\gamma_B)^2}\left(\sqrt{\frac{(1-\pi)\gamma_A+\pi\gamma_B\rho_{\text{FULL}}^A}{(1-\pi)\gamma_A+\pi\gamma_B}}\right)^{-1},
$$

and is positive since $\rho_{\text{FULL}}^A<1$. A similar argument applies for group $g=B$ since $\rho_{\text{FULL}}^B>1$.

Step 3: We show that for any given group g and $\gamma_{g'}\in(0,1]$, $g'\neq q$, there exists threshold $\bar{\bar{\gamma}}_g\in(0,1]$ such that academic merit of group g improves if and only if $\gamma_g<\bar{\bar{\gamma}}_g$. Fix group A; the proof is analogous for group B. It suffices to show that (a) $\bar{\bar{\gamma}}_A$ is the unique solution to $\beta_A(\bar{\bar{\gamma}}_A,\gamma_B,\rho_{\text{FULL}}^A)=\Delta_A(\xi_A,\rho_{\text{SUB}}^A)$ and (b) $\bar{\bar{\gamma}}_A\in(0,1]$.

Conditional on the existence of $\bar{\bar{\gamma}}_A$, uniqueness in (a) follows immediately from the monotonicity of β_A shown in Step 2. Existence in turn can be shown as follows. In the absence of barriers part (iii)

in Theorem 2 guarantees that the academic merit of group g decreases after dropping test scores, thus $\beta_A(1, \gamma_B, \rho_{\text{FULL}}^A) > \Delta_A(\xi_A, \rho_{\text{FULL}}^A)$. Furthermore, observe that for $\gamma_A = 0$, academic merit trivially improves from $\beta_A(0, \gamma_B, \rho_{\text{FULL}}^A) = 0$ to a positive value $\Delta_A(\xi_A, \rho_{\text{SUB}}^A) > 0$ after dropping test scores. Thus, by the continuity of $\beta_A(\gamma_A, \gamma_B, \rho_{\text{FULL}}^A)$, such a $\bar{\bar{\gamma}}_A$ exists. For part (b), continuity of β_A further implies that there must exist an interval $[0, \epsilon)$, $\epsilon > 0$, such that $\beta_A(\gamma_A, \gamma_B, \rho_{\text{FULL}}^B) < \Delta_A(\xi_A, \rho_{\text{SUB}}^A)$ for all $\gamma_A \in [0, \epsilon)$. Consequently, $\bar{\bar{\gamma}}_A \geq \epsilon > 0$. $\qquad\square$

Proof of part (ii). Plugging (20) into the definition of diversity with and without test scores, respectively, it immediately follows that diversity improves if and only if $\eta(1, 1, \rho_{\text{SUB}}^B) > \eta(\gamma_A, \gamma_B, \rho_{\text{FULL}}^B)$.

Step 1: Fix all parameters (including $\gamma_A \in (0, 1]$) except for $\gamma_B \in (0, 1]$. We show that diversity strictly increases as barriers decrease (γ_B increases), i.e., $\eta(\gamma_A, \gamma_B', \rho_{\text{FULL}}^B) > \eta(\gamma_A, \gamma_B, \rho_{\text{FULL}}^B)$ for $\gamma_B' > \gamma_B$.

By (20), the admission threshold increases as γ_B increases. Indeed, $\tilde{q}_{\text{SUB}}^*$ is the solution to $(1 - \pi)\gamma_A(1 - F_{\tilde{q}|A, P_{\text{SUB}}}(\tilde{q}_{\text{SUB}}^*)) + \pi\gamma_B(1 - F_{\tilde{q}|B, P_{\text{SUB}}}(\tilde{q}_{\text{SUB}}^*)) = 1 - C$. Thus, as γ_B increases, the solution $\tilde{q}_{\text{SUB}}^*$ must decrease since each $F_{\tilde{q}|g, P_{\text{SUB}}}$ is increasing in its argument.

Then, since the admission threshold $\tilde{q}_{\text{SUB}}^*$ increases but the capacity C, barriers γ_A (thus the mass of students in group A who are eligible to apply), and the perceived skill distributions for both groups remain constant, it follows that a lower mass of students are admitted from group A. As a result, the remaining capacity is filled with more students from group B, which in turn implies that diversity increases.

Step 2: We show that, given all other parameters fixed including γ_A, there exists a threshold $\bar{\gamma}_B(\gamma_A)$ such that diversity increases after dropping the test if and only if $\gamma_B < \bar{\gamma}$. It suffices to show that (a) $\bar{\gamma}$ is the unique solution to $\eta(1, 1, \rho_{\text{SUB}}^B) = \eta(\gamma_A, \bar{\gamma}, \rho_{\text{FULL}}^B)$ and (b) $\bar{\gamma} \in (0, 1]$. The proof follows as in Step 3 in Part (i). $\qquad\square$

C.6 Affirmative action

We examine policies P_S^τ with *affirmative action*, meaning that the school sets a target diversity level $\tau(P_S^\tau) = \tau$. Thus, the common threshold $\tilde{q}_S^*$ in (10) is replaced by two group-dependent thresholds, $q_{A,S}^*$ and $q_{B,S}^*$:

$$(1 - \pi)\gamma_A(1 - F_{\tilde{q}|A, P_S}(q_{A,S}^*)) = (1 - \tau)C, \quad \pi\gamma_B(1 - F_{\tilde{q}|B, P_S}(q_{B,S}^*)) = \tau C. \tag{22}$$

Note further that the distribution $F_{\tilde{q}|g, P_S} \equiv F_{\tilde{q}|g, P_S^\tau}$, $g \in \{A, B\}$, remains unchanged under both admissions policies P_S^τ and P_S, as both share the same group-aware estimation policy and feature set S.

Proposition 2 (Affirmative action with a fixed testing policy)**.** *Fix the target diversity level $\tau(P_S) < \tau \leq \pi$ and assume unequal precisions. Let also $\gamma_B \leq \gamma_A \leq 1$ such that $\gamma_A \geq \frac{2(1-\tau)C}{1-\pi}$, $\gamma_B \geq \frac{2\tau C}{\pi}$. Then,*

(i) Individual fairness: *In comparison to P_S, the individual fairness gap improves, i.e., $I(q; P_S^\tau) < I(q; P_S)$ for all q. However, group A students still have higher probability of admission than same-skilled group B students, i.e., $I(q; P_S) > 0$, if and only if*

$$q > \frac{\left(\frac{\sum_{k \in S} \sigma_{Ak}^{-2} + \sigma^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}\right) \tilde{q}_{A,S}^* - \left(\frac{\sum_{k \in S} \sigma_{Bk}^{-2} + \sigma^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}\right) \tilde{q}_{B,S}^*}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} - \sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}} + \frac{\mu \sigma^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} \sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}.$$

Finally, there exist parameters such that $I(q; P_S^\tau) < 0 < I(q; P_S)$ for some q.

(ii) Academic merit: *Policy P_S^τ always achieves worse academic merit for admitted group B students than for group A students. Furthermore, in comparison to P_S, the academic merit of admitted students decreases for group B, while it increases for group A.*

Proof of Part (i). First, note that (22) gives us

$$\tilde{q}_{A,S}^* = F_{\tilde{q}|A,P_S}^{-1} \left(1 - \frac{1 - \tau}{(1 - \pi)\gamma_A} C\right), \quad \tilde{q}_{B,S}^* = F_{\tilde{q}|B,P_S}^{-1} \left(1 - \frac{\tau}{\pi \gamma_B} C\right). \tag{23}$$

Since $\tau > \tau(P_S)$ and $\gamma_B \leq \gamma_A \leq 1$, it follows that $\tilde{q}_{B,S}^* < \tilde{q}_S^* < \tilde{q}_{A,S}^*$. Due to our assumptions that $\gamma_A \geq \frac{2(1-\tau)C}{1-\pi}$ and $\gamma_B \geq \frac{2\tau C}{\pi}$, we also get that $\mu < \tilde{q}_{B,S}^* < \tilde{q}_S^* < \tilde{q}_{A,S}^*$.

For the first statement of part (i), observe that, due to $\tilde{q}_{A,S}^* > \tilde{q}_S^*$ and $\tilde{q}_{B,S}^* < \tilde{q}_S^*$ for all $\tau(P_S) < \tau \leq \pi$, $\mathbb{P}[\tilde{q} \geq \tilde{q}_{A,S}^* \mid q, A, P_S^\tau] < \mathbb{P}[\tilde{q} \geq \tilde{q}_S^* \mid q, A, P_S]$, and $\mathbb{P}[\tilde{q} \geq \tilde{q}_{B,S}^* \mid q, B, P_S^\tau] > \mathbb{P}[\tilde{q} \geq \tilde{q}_S^* \mid q, B, P_S]$, since the distribution of $\tilde{q} \mid q, P$ remains the same under both $P \in \{P_S, P_S^\tau\}$. Consequently, $I(q; P_S^\tau) < I(q; P_S)$.

For the proof of the second statement in Part (i), we apply the argument used in Proposition 1, Part (ii). Thus, we get that $I(q; P_S^\tau) > 0$ if and only if

$$\frac{\tilde{q}_{A,S}^* \sigma^{-2} + \tilde{q}_{A,S}^* \sum_{k \in S} \sigma_{Ak}^{-2} - \mu \sigma^{-2} - q \sum_{k \in S} \sigma_{Ak}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}} < \frac{\tilde{q}_{B,S}^* \sigma^{-2} + \tilde{q}_{B,S}^* \sum_{k \in S} \sigma_{Bk}^{-2} - \mu \sigma^{-2} - q \sum_{k \in S} \sigma_{Bk}^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}},$$

which is equivalent to

$$q > \frac{\left(\frac{\sum_{k \in S} \sigma_{Ak}^{-2} + \sigma^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}}}\right) \tilde{q}_{A,S}^* - \left(\frac{\sum_{k \in S} \sigma_{Bk}^{-2} + \sigma^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}\right) \tilde{q}_{B,S}^*}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} - \sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}} + \frac{\mu \sigma^{-2}}{\sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} \sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}}}.$$

Finally, we prove the third statement in Part (ii). Consider an instance Ω where

$$\sqrt{\sum_{k \in S} \sigma_{Bk}^{-2}} \sqrt{\sum_{k \in S} \sigma_{Ak}^{-2}} > \sigma^{-2}, \tag{24}$$

and under P_S^τ, the condition in Part (ii) in Proposition 1, holds with equality for some $\hat{q}$, i.e.,

$$(\tilde{q}_{A,S}^* - \hat{q})\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} = \sigma^{-2}(\tilde{q}_{A,S}^* - \mu).$$

Therefore, $\mathbb{P}[\tilde{q} > \tilde{q}_{A,S}^* \mid \hat{q}, A] = \mathbb{P}[\tilde{q} > \tilde{q}_{A,S}^* \mid \hat{q}, B]$. Since $\tilde{q}_{B,S}^* < \tilde{q}_{A,S}^*$, it further holds that $\mathbb{P}[\tilde{q} > \tilde{q}_{B,S}^* \mid \hat{q}, B] > \mathbb{P}[\tilde{q} > \tilde{q}_{A,S}^* \mid \hat{q}, B]$. Thus, $I(\hat{q}; P_S^\tau) < 0$.

However, for $q = \hat{q}$, we also have that

$$(\tilde{q}_S^* - \hat{q})\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} < \sigma^{-2}(\tilde{q}_S^* - \mu).$$

To see why, observe that given the condition in (24), the function

$$g(\tilde{q}) = (\tilde{q} - \hat{q})\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} - \sigma^{-2}(\tilde{q} - \mu)$$

is increasing in $\tilde{q}$ since

$$\frac{\mathrm{d}g(\tilde{q})}{\mathrm{d}\tilde{q}} = \sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}} - \sigma^{-2} > 0.$$

Consequently, for $\tilde{q}_S^* < \tilde{q}_{A,S}^*$, $g(\tilde{q}_S^*) < g(\tilde{q}_{A,S}^*) = 0$. Part (ii) in Proposition 1 further guarantees that $I(\hat{q}; P_S) > 0$ for instance Ω. Finally, we have constructed a problem instance Ω such that $I(\hat{q}; P_S) > 0 > I(\hat{q}; P_S^\tau)$ for some $\hat{q}$. Thus, such an instance exists. $\qquad\square$

Proof of Part (ii). We use an argument similar to part (iii) in Proposition 1 (note that this part holds for any common threshold greater than μ and not only $\tilde{q}_S^*$). Similarly to (16), we derive that for both $g \in \{A, B\}$, $\mathbb{E}[q \mid \tilde{q} \geq \tilde{q}_{g,S}^*, g, P_S^\tau] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{g,S}^*, g, P_S^\tau]$. By the same part (iii) in Proposition 1, replacing $\tilde{q}_S^*$ with threshold $\tilde{q}_{A,S}^* > \mu$ implies that $\mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{A,S}^*, A, P_S^\tau] > \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{A,S}^*, B, P_S^\tau]$. Next, we have that

$$\begin{aligned}
\mathbb{E}[\tilde{q} \mid Y = 1, B, P_S^\tau] &= \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{B,S}^*, B, P_S^\tau]\\
&= \frac{1}{1 - F_{\tilde{q}|B,P_S}(\tilde{q}_{B,S}^*)}\int_{\tilde{q}_{B,S}^*}^{\infty}\tilde{q}\,dF_{\tilde{q}|B,P_S}(\tilde{q})\\
&= \frac{1}{1 - F_{\tilde{q}|B,P_S}(\tilde{q}_{B,S}^*)}\left(\int_{\tilde{q}_{B,S}^*}^{\tilde{q}_{A,S}^*}\tilde{q}\,dF_{\tilde{q}|B,P_S}(\tilde{q}) + \int_{\tilde{q}_{A,S}^*}^{\infty}\tilde{q}\,dF_{\tilde{q}|B,P_S}(\tilde{q})\right)\\
&= \frac{F_{\tilde{q}|B,P_S}(\tilde{q}_{A,S}^*) - F_{\tilde{q}|B,P_S}(\tilde{q}_{B,S}^*)}{1 - F_{\tilde{q}|B,P_S}(\tilde{q}_{B,S}^*)}\mathbb{E}[\tilde{q} \mid \tilde{q}_{A,S}^* > \tilde{q} \geq \tilde{q}_{B,S}^*, B, P_{a,K}^\tau]\\
&\quad + \frac{1 - F_{\tilde{q}|B,P_S}(\tilde{q}_{A,S}^*)}{1 - F_{\tilde{q}|B,P_S}(\tilde{q}_{B,S}^*)}\mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{A,S}^*, B, P_{a,K}^\tau].
\end{aligned}$$

The fact that $\mathbb{E}[\tilde{q} \mid \tilde{q}_{A,S}^* > \tilde{q} \geq \tilde{q}_{B,S}^*, B, P_S^\tau] < \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{A,S}^*, B, P_S^\tau]$, together with the

inequalities above, finally imply that

$$\mathbb{E}[q \mid Y = 1, B, P_S^\tau] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{B,S}^*, B, P_S^\tau] < \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{A,S}^*, A, P_S^\tau] = \mathbb{E}[q \mid Y = 1, A, P_S^\tau].$$

Regarding the second statement of part (ii), recall that the distributions $F_{\tilde{q}|g,P_S}$ and $F_{\tilde{q}|g,P_S^\tau}$ are identical. Since $\tilde{q}_{B,S}^* < \tilde{q}_S^* < \tilde{q}_{A,S}^*$, it follows that the conditional expectations satisfy

$$\mathbb{E}[q \mid Y = 1, A, P_S^\tau] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{A,S}^*, A, P_S^\tau] > \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_S^*, A, P_S] = \mathbb{E}[q \mid Y = 1, A, P_S],$$

$$\mathbb{E}[q \mid Y = 1, B, P_S^\tau] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_{B,S}^*, B, P_S^\tau] < \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_S^*, B, P_S] = \mathbb{E}[q \mid Y = 1, B, P_S].$$

Thus, the academic merit of admitted students increases for group A while it decreases for group B. $\qquad\square$

Tension between barriers and informativeness (with affirmative action). We study how test-free and test-based policies with affirmative actions compare to each other in a setting with barriers to testing.

Proposition 3 (Dropping tests under affirmative action with barriers). *Fix group $g \in \{A, B\}$, variances σ_{gk}^2, and target diversity level τ. Let $\tau_A \triangleq 1 - \tau$ and $\tau_B \triangleq \tau$. Dropping the test score requirement improves the academic merit of admitted students from group g, i.e., $\mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}^\tau] < \mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}^\tau]$, if and only if $\gamma_g \leq \hat{\gamma}_g$, where*

$$\hat{\gamma}_g = \cfrac{\tau_g C}{1 - \Phi\left(\mathrm{HR}^{-1}\left(\cfrac{\sqrt{\cfrac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}}}{\sqrt{\cfrac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}}} \mathrm{HR}(\Phi^{-1}(1 - \tfrac{\tau_g C}{\pi_g}))\right)\right)}. \tag{5}$$

Fixing all other parameters, the threshold $\hat{\gamma}_g$ increases as test variance σ_{gK} for group g increases.

Proof. Let $\tilde{w}_{g,\text{FULL}}^*$ be the group-dependent threshold in a policy with barriers and affirmative action. Define

$$t_g = \frac{\tilde{w}_{g,\text{FULL}}^* - \mu}{\sigma \sqrt{\cfrac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}}}, \quad t_g' = \frac{\tilde{q}_{g,\text{SUB}}^* - \mu}{\sigma \sqrt{\cfrac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}}}.$$

For such a policy with admission thresholds $\tilde{w}_{g,\text{FULL}}^*$, $g \in \{A, B\}$, Lemma C.3 implies that the expected skill level of admitted students in group g equals

$$\mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}^\tau] = \mu + \sigma \sqrt{\cfrac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}} \cdot \frac{\phi(t_g)}{1 - \Phi(t_g)}.$$

Similarly, for a policy using affirmative action but no tests, and admission thresholds $\tilde{q}_{g,\text{SUB}}^*$, we get

that

$$\mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}^\tau] = \mu + \sigma \sqrt{\frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}} \cdot \frac{\phi(t_g')}{1 - \Phi(t_g')}.$$

To compute the threshold $\hat{\gamma}_g$, we require that $\mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}^\tau] = \mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}^\tau]$. Based on the above equations, this condition is equivalent to

$$\sqrt{\frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}} \text{HR} \left(\frac{\tilde{q}_{g,\text{SUB}}^* - \mu}{\sigma \sqrt{\frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}}} \right) = \sqrt{\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}} \text{HR} \left(\frac{\tilde{w}_{g,\text{FULL}}^* - \mu}{\sigma \sqrt{\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}}} \right).$$

Letting $\tau_B = \tau$, $\tau_A = 1 - \tau$ and using (23) to compute the thresholds $\tilde{w}_{g,\text{FULL}}^*$, $\tilde{q}_{g,\text{SUB}}^*$, we get that

$$\sqrt{\frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}} \text{HR} \left(\Phi^{-1} \left(1 - \frac{\tau_g C}{\pi_g} \right) \right) = \sqrt{\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}} \text{HR} \left(\Phi^{-1} \left(1 - \frac{\tau_g C}{\pi_g \hat{\gamma}_g} \right) \right)$$

Thus, solving for $\hat{\gamma}_g$, we finally get (5). Note that the expected skill level of admitted students in the test-based policy is given – due to Lemma C.3 – by

$$\mu + \sigma \sqrt{\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}} \text{HR} \left(\Phi^{-1} \left(1 - \frac{\tau_g C}{\pi_g \hat{\gamma}_g} \right) \right).$$

By Lemma C.4, it follows that HR is increasing. However, $\Phi^{-1} \left(1 - \frac{\tau_g C}{\pi_g \hat{\gamma}_g} \right)$ is decreasing in $\hat{\gamma}_g$. Therefore, the academic merit of g must be decreasing in $\hat{\gamma}_g$. Thus, dropping the test increases academic merit for g if and only if $\gamma_g \leq \hat{\gamma}_g$.

Finally, we prove the second claim. As σ_{gK} increases, $\sqrt{\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}}$ decreases. Thus, the quantity

$$\frac{\sqrt{\frac{\sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}}}{\sqrt{\frac{\sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}}} \text{HR} \left(\Phi^{-1} \left(1 - \frac{\tau_g C}{\pi_g} \right) \right)$$

increases. By Lemma C.4, the hazard rare (HR) is increasing so its inverse HR^{-1} is also increasing. Since the CDF Φ is increasing, their composition $\Phi(\text{HR}^{-1}(\cdot))$ must be also increasing, which in turn implies that the denominator in (5) is decreasing in σ_{gK}. Consequently, $\hat{\gamma}_g$ increases as σ_{gK} increases. $\square$