Electronic Companion: *Dropping Standardized Testing for Admissions Trades Off Information and Access* (by Nikhil Garg, Hannah Li, and Faidra Monachou)

EC.1. Calibrated simulations with UT Austin data

We calibrate our model to empirical data to assess the effects of dropping test requirements under our model. Our results establish that (a) there are reasonable parameter ranges both in which dropping the test can be beneficial and harmful for the desiderata, and (b) when tests are required, outcomes can depend on whether the model allows students to self-select to take the test. Real admissions decisions are much more complex than our model (Selingo 2020), and a key challenge in empirical admissions settings is selective data availability (Rothstein 2004), since we typically only observe college outcome data for those admitted, which partially depends on test scores and other admission features. Given these limitations, our calibrated simulation exercise should be viewed as suggestive examples that dropping the test can either improve or worsen the desiderata, as opposed to establishing optimal policy for any particular setting.

Data. Our data is from the Texas Higher Education Opportunity Project (THEOP), a semipublic 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 enrolled there in 1992-1997 and completed at least 24 credit hours.[12] For each student, 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 characteristics of their high school (including relative *economic privilege* rounded to nearest quartile, which is a measure of the socioeconomic status of the students the high school serves). Since we consider enrolled students, we observe their GPA and number of credit hours for each enrolled semester, that we use to calculate overall GPA in their first year and afterwards.

Calibration and simulation setup. We conduct a calibrated simulation exercise for a hypothetical admissions setting in which the applicant population looks distributionally similar to students who in reality enrolled to UT Austin.[13] For each individual, we use their

[12] This period represents admissions 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.

[13] This could reflect, for example, admissions at a college more selective than UT Austin in the time period considered.

cumulative *college GPA*—not counting their first year—to represent their true skill. Then, as features, we use (in various simulations) their high school *class rank*, standardized *test score* and/or college *first-year GPA*. To form the two groups, we take the upper (group A) and lower (group B) halves of the high schools' *economic privilege*[14] index.

We calibrate our model parameters to the empirical data. We calibrate the true skill mean μ and variance σ^2 to the empirical mean and variance of the cumulative college GPA, excluding the first year. We then calibrate the conditional feature distributions for each group, which in our model are distributed as $\theta_k \sim N(q+\mu_{gk}, \sigma_{gk}^2)$; i.e., for each group g and feature k pair, we need estimates of μ_{gk} and σ_{gk}^2, the *conditional* mean and variance of the feature given the student's true skill. We estimate these values by running an ordinary least squares regression $\theta_k \sim q$, where q is the observed college GPA. Let the fitted regression model be $\hat{\theta}_k = \hat{\beta}_0 + \hat{\beta}_1 q$, so that $q = (\hat{\theta}_k - \hat{\beta}_0)/\hat{\beta}_1$. To normalize the features so that a one unit increase in the feature corresponds to a one unit increase in skill level (so that the feature has mean $q+\mu_{gk}$), we center and scale each observed feature to obtain $\theta'_k = (\theta_k - \hat{\beta}_0)/\hat{\beta}_1$, and likewise for the predicted features $\hat{\theta}_k$ to obtain $\hat{\theta}'_k$. Now, we calibrate the model to the distribution of θ'_k. We set μ_{gk} to be the sample mean of θ'_k and σ_{gk}^2 to be the sample variance of the residuals $\hat{\theta}'_k - \theta'_k$. The calibrated standard deviation parameters σ_{gk} are in Table EC.1.

Group	HS class rank	College GPA, 1st year	Test score
A (high economic privilege)	2.00	0.98	3.30
B (low economic privilege)	2.65	0.91	3.11

Table EC.1 Calibrated feature standard deviations σ_{gk} for each group g and feature θ_k. This calibration suggests that class rank is relatively more predictive of cumulative college GPA for the high economic privilege group while the test score is more predictive for the low economic privilege group – consistent with UC Standardized Testing Task Force (2020) and Schmill (2022). Most predictive for each is the first-year college GPA.

Using these calibrated mean and variance parameters, we then simulate our model, with the students' applications and the school's Bayesian updating as described earlier. We simulate the admission outcomes in both the setting with strategic students and the

[14] Column by the data provider, defined as "Publicly available data from the Texas Education Agency (TEA) is used to stratify regular, Texas public high schools according to the socioeconomic status of the students they serve. The 25% of high schools with the lowest percent of students ever economically disadvantaged are designated as Upper quartile. The 25% of high schools having the highest percent of students ever economically disadvantaged are designated as Lower quartile. Because the statewide share of economically disadvantaged students rose over time, quartile cut points are calculated separately for each year." We then binarize the quartiles.

setting with non-strategic students. In both settings, we fix group A to have full access to the test ($\gamma_A = 1$ and $c_A = 0$ in the non-strategic and strategic settings, respectively) and vary the level of access for group B students. We set the student utility for the school to be $v = 5$. We fix an equal proportion of students from each group in the candidate pool ($\pi = 0.5$). We simulate a setting with 10,000 applicants and a capacity of 1,000. For each parameter set, we run 100 simulations and report the mean and 95% confidence intervals across simulation runs.

We simulate two informational cases, which correspond to the school having access to different features when making its decision.

Low informativeness: Class rank and (potentially) *Test score*. Simulates, for example, a setting in which the application pool and information available is incoming first-year students at UT Austin.

High informativeness: First-year GPA and (potentially) *Test score*. Simulates, for example, a setting in which the application pool and information available is students at the end of their first year at UT Austin.

To make the non-strategic and strategic settings comparable, we define the notion of *test access level* for group B as the proportion of group B students taking the test. In the non-strategic setting, this is γ_B by definition. In the strategic setting, each cost level c_B induces a test access level which can be found through simulation. We note that while the overall number of group B students taking the test is the same for a fixed test access level, in the strategic setting this group of students are disproportionately high-skilled (see Lemma 3 and Figure EC.4).

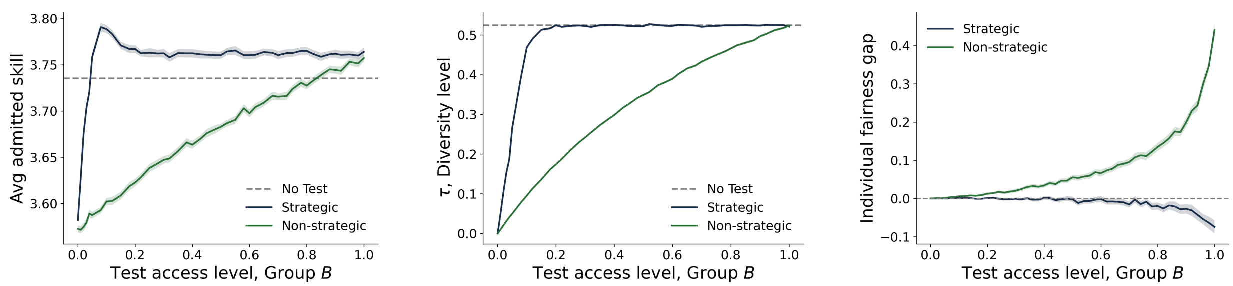
Simulation results. Table EC.2 summarizes the admission outcomes with and without the test, for a fixed level of group B students having access (40%), while all group A students have access.[15] For outcomes for the full range of group B test access, see Figures EC.1 and EC.2, for the high and low informational environment, respectively.

In this setting, exactly half of the students are in group B ($\pi = 0.5$). For any diversity level below 50% (i.e., students in group B make up less than half of the admitted student body), we consider group B to be under-represented.

[15] Using the College Board (2022) California SAT Suite of Assessments Annual Report, we calculate that a student from the bottom two quintiles of family income are 38% as likely to take the test as a student from the top two quintiles. Thus we focus on an access levels of 100% and 40% for groups A and B, respectively.

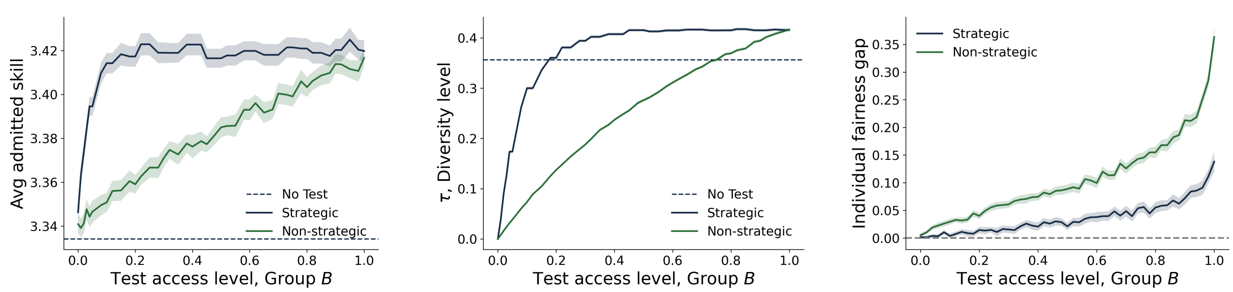
		Academic merit		Diversity Level	
Informational Case	**Student behavior**	With test	Without test	With test	Without test
Low	Strategic	3.42	3.33	40.8%	35.7%
		± .005	± .0005	± .3%	± .03 %
	Non-strategic	3.38	3.33	23.7%	35.7%
		± .005	± .0005	± .3%	± .03%
High	Strategic	3.76	3.74	52.5%	52.4%
		± .005	± .0004	± .3%	± .03%
	Non-strategic	3.66	3.74	29.4%	52.4%
		± .004	± .0004	±.3%	± .03 %

Table EC.2 How academic merit and diversity level of admitted students change with and without requiring a test score, for two informational cases (how informative the non-test feature is) and for the strategic and non-strategic settings. Academic merit (GPA) ranges from 1.0-4.0. Diversity level is shown as a percentage of admitted students. Shown with 95% confidence intervals. This table assumes a 40% test access for group B; for outcomes for the full range of group B test access, see Figures EC.1 and EC.2. Values are averaged across 100 simulation runs with 95% confidence intervals shown. All differences are statistically significant.



(a) Academic merit (b) Diversity level (c) Individual fairness gap

Figure EC.1 Calibrated simulations when full set of features is {First year GPA, test score} – high informativeness case (see Table EC.1 for informativeness of features). Figures (a) and (b) vary test the access level for group B. Figure (c) shows the individual fairness gap at a fixed group B test access level of 40%. Figures show average value across 100 simulation runs and 95% confidence intervals.



(a) Academic merit (b) Diversity level (c) Individual fairness gap

Figure EC.2 Calibrated simulations when full set of features is {High school class rank, test score}. Low informativeness case. See Table EC.1 for informativeness of features.

Overall, the results show that the effects of dropping the test requirement depend crucially on both the informational environment and whether students are strategic. At a test access level of 40% for group B, dropping the test worsens both academic merit and diversity level when students are strategic in both informational cases, although only slightly in the high information case. However, when students are non-strategic, dropping the test improves both metrics when the remaining feature has high informativeness, whereas dropping the test has mixed results when the other feature has low informativeness.

Comparing effect of test access in strategic and non-strategic settings. The results show that in both informational settings, academic merit, diversity, and individual fairness all worsen when fewer group B students have access to the test. However, for a given level of test access, the outcomes for all three metrics are better when students are strategic, compared to when they are non-strategic. In the strategic setting, the students with higher skill levels are more likely to take the test (see Lemma 3 and Figure EC.4), as opposed to the non-strategic setting where all students in group B have the same probability γ_B of taking the test. Thus, as we see in Figures EC.1 and EC.2, even when the test access levels are as low as 30 percent, the admission outcomes of academic merit, diversity, and individual fairness are comparable to when group B has full test access. This observation, of course, relies on the students appropriately assessing their likelihood of admission upon taking the test, which we assume in our model. We also note that academic merit in particular is not monotonic in the test access level (Figure EC.1). As the access level for group B approaches 0, the average skill level for admitted students increases for group B but decreases for group A, leading to non-monotonicity in the overall academic merit. See Figure EC.6b for an illustration of average skill level of admitted students, by group.

Effect of 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 and diversity when costs are high enough; 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, thus leading to worse admissions outcomes.

These findings underscore our theoretical results: the consequences of dropping test scores depend crucially on the information content of other signals, the level of strategic behavior by applicants, and the levels of access to the test. Decisions to require the test should not (and cannot) be made in a context-independent manner.

Discussion. There are several ways in which our simulation setup differs from reality, for example: (1) We use college GPA as a measure of student true skill; in reality, GPA is a function of many other aspects as well, such as college major and barriers faced during college (Engle and Tinto 2008). (2) Because of our choice to use college GPA as a true skill measure, we cannot simulate our model for all students who *apply* to UT Austin, as data is censored[16]—we do not observe their college GPA unless they enrolled. Thus, we must simulate a hypothetical admissions setting for which the enrolled population at UT Austin is a reasonable application pool. (3) To closely simulate our model, we fit Normal distributions to the data, while the respective distributions may not be Normally distributed (e.g., many of the features are truncated). (4) We do not have estimates of the barriers or costs to testing, and in fact almost all applicants in the data (over 99.9%) have test scores due to school policies at the time; thus, we have to artificially simulate some students as not having access. For these reasons, our simulations should not be interpreted as making statements about the UT Austin context or any particular admissions setting.

EC.2. Simulations with synthetic data

EC.2.1. Supplementary simulations for the non-strategic setting

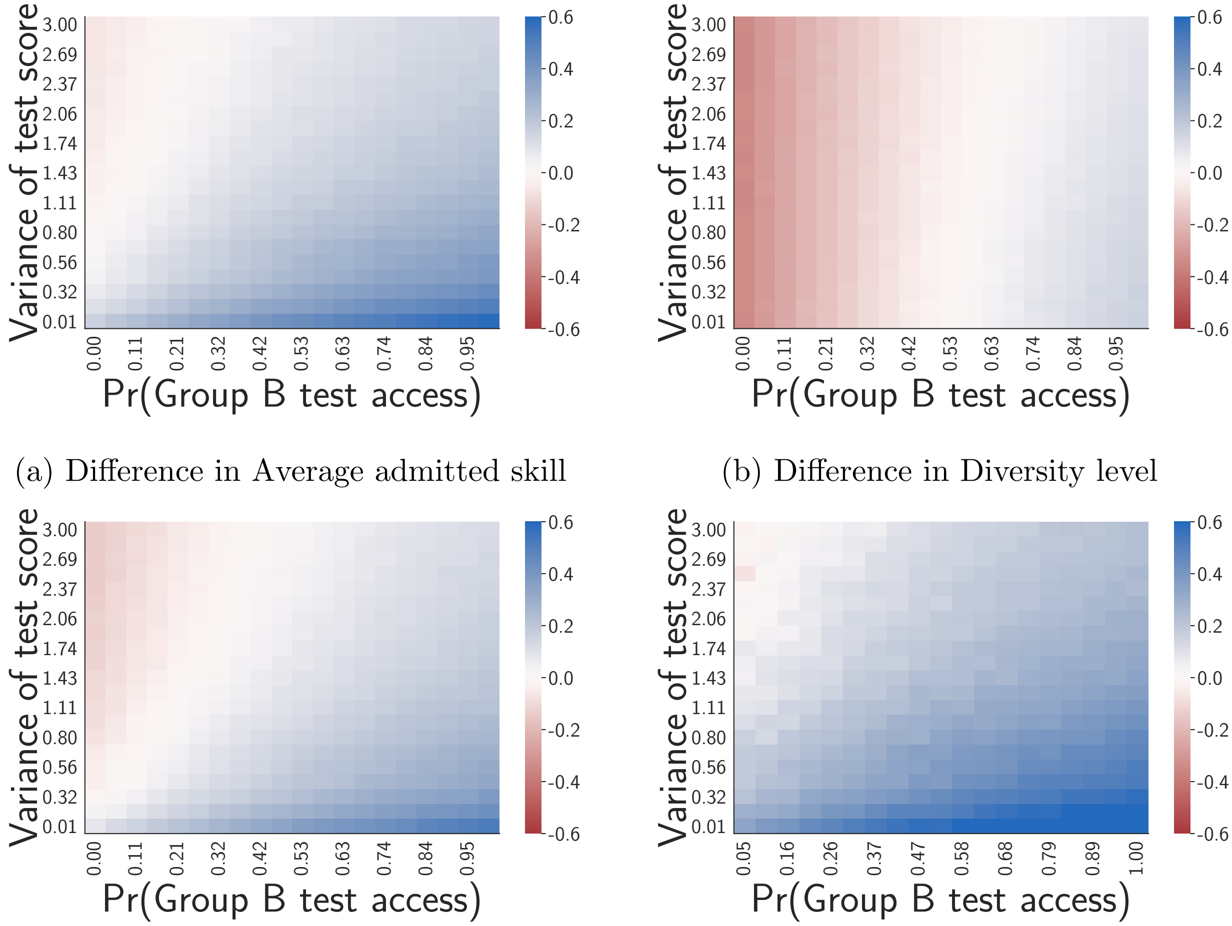
EC.2.1.1. Supplemental simulation figures for the non-strategic setting Figure EC.3 supplements the results in Theorem 1 and Proposition EC.1, regarding the thresholds at which academic merit and diversity improve after dropping the test. In particular, they illustrate that for high enough test score variance or high enough barriers, dropping the test score improves the objectives.

EC.2.1.2. Simulation parameters We report the parameters for the simulations with non-stretegic students.

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 $\theta_{B1} \sim N(-4, \sigma^2_{B1})$, where $\sigma^2_{B1} \in (0,5)$. For subfigures (3a) - (3c) we fix test access $\gamma_A = \gamma_B = 1$. For subfigures (3d) - (3f) fix the test score variance of group B to be equal to that of group A, so that $\sigma^2_{B1} = \sigma^2_{A1} = 1$ and we vary γ_B.

[16] This is a common barrier to measuring the predictive power of standardized testing in admissions (Rothstein 2004, Weissman 2020).



(a) Difference in Average admitted skill

(b) Difference in Diversity level

(c) Difference in Average admitted skill, Group A

(d) Difference in Average admitted skill, Group B

Figure EC.3 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 Electronic Companion EC.2.1.2. The plot reads as follows: in Figure EC.3a, 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.

Figure EC.3. 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, $\gamma_B \in (0,1)$.

Figure EC.9. Same as Figure 2.

EC.2.2. Additional simulations for the strategic setting

In this appendix, we report results for the setting with strategic students. Section EC.2.2.1 focuses on a single school and Section EC.2.3 studies two competing schools.

EC.2.2.1. Simulations with synthetic data and a single school We first describe the general simulation setup for a single school and then report the parameters and other details for each figure.

Simulation setup for strategic students and a single school. Fix a single school that uses policy P_{FULL}. All students are initialized with realizations of their true skill q and non-test features $\boldsymbol{\theta}_{\text{SUB}}$. Students follow the behavior outlined in the proof of Lemma 3 (see Appendix EC.4.5). Note that this result shows the existence of an equilibrium that is characterized by thresholds $\underline{q}^g_{\text{SUB}}$, where students take the test if and only if $\tilde{q}_{\text{SUB}} \geq \underline{q}^g_{\text{SUB}}$ but does not directly give the value of these thresholds. We simulate admissions process under candidate equilibrium thresholds $\underline{q}^g_{\text{SUB}}$, each of which results in a certain number of students being admitted. We find the thresholds $\underline{q}^g_{\text{SUB}}$ such that the school's capacity constraint is respected and size of the admitted student body is the closest to the threshold.

The school's admission decisions and students' test taking decisions follow the proof of Lemma 3. For the policy P_{FULL}, fix a candidate admission threshold $\tilde{q}'$ for the estimated skill $\tilde{q}_{\text{FULL}}|\boldsymbol{\theta}_{\text{FULL}}, g$. First, consider the students' observations. Each student observes their non-test features θ_{SUB} and estimates the distribution of their estimated skill if they were to take the test $\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}$, given in Equation (EC.20). The student then calculate $\mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) = \mathbb{P}(\tilde{q}(\boldsymbol{\theta}_{\text{FULL}}, g) \geq \tilde{q}^*_{\text{FULL}} \mid \boldsymbol{\theta}_{\text{SUB}}, g)$ and solves for their optimal test taking decision $\arg\max_{\alpha\in\{0,1\}} \alpha\,(v\,\mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) - c_g)$, as seen in Lemma EC.11. In other words, the student takes the test when $v\,\mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) \geq c_g$. The students who take the test then apply to the school. Now, the school admits all students with estimated skill $\tilde{q}_{\text{FULL}}|\boldsymbol{\theta}_{\text{FULL}}, g \geq \tilde{q}'$. Note, however, that this may result in a smaller or larger admitted class than the school's capacity. We then search across candidate thresholds $\tilde{q}'$ and set $\tilde{q}^*_{\text{FULL}}$ to be the q' that attains the largest admitted class size, while still respecting the capacity constraint.

Parameters for Figure EC.4. Figure EC.4 shows simulation results illustrating the student equilibrium decisions (characterized by Equation (5)) on whether to take the test and apply to a school that requires the test, as a function of their true skill and group.

There are two features, where the where the non-test feature is equally informative for both groups, but the test score is more informative for group A than group B. The true skill distribution for both groups is Normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 1$. The features for the two groups are Normally distributed with mean $\mu_{gk} = 0$ for all

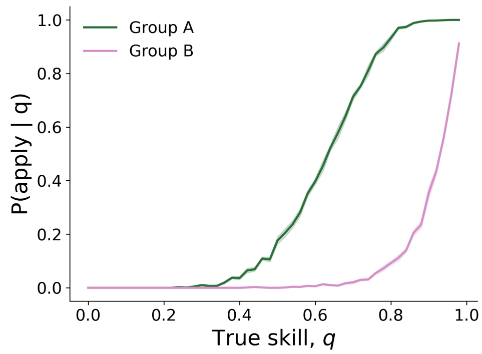
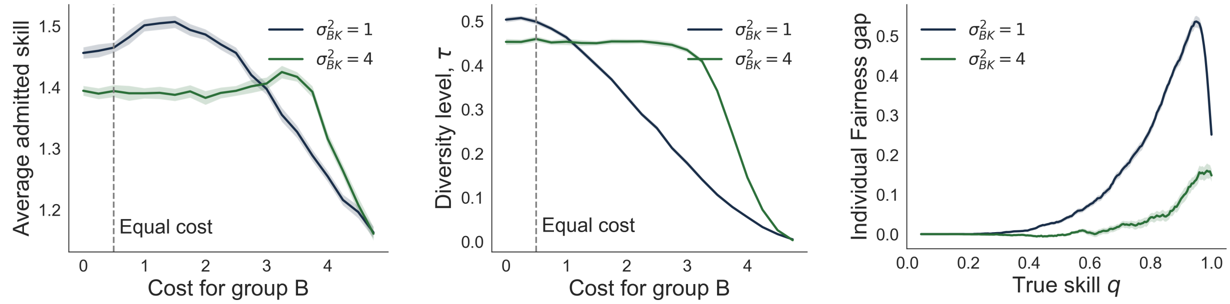


Figure EC.4 Results of a simulation calculating the student equilibrium decisions (characterized by Equation (5)) on whether to take the test and apply to a school that requires the test, as a function of their true skill and group. When students are strategic, high skilled students are more likely to take the test and apply. We consider a case where Group A has higher precision and lower test costs: Parameters are $\mu_{gk} = 0$ for all g, k, $\sigma^2_{A1} = \sigma^2_{A2} = \sigma^2_{B1} = 1$ and $\sigma^2_{B2} = 2$, where $k = 2$ denotes the test feature. Costs are $c_A = 0.5$ and $c_B = 3$. In this case, Group A students are more likely to take the test than Group B students of the same true skill value. Full simulation parameter set can be found in Electronic Companion EC.2.2.1.

g, k and $\sigma^2_{A1} = \sigma^2_{A2} = \sigma^2_{B1} = 1$ and $\sigma^2_{B2} = 2$, where $k = 2$ denotes the test feature. Students of both groups have valuation $v = 5$ for the school. Test costs are $c_A = 0.5$ and $c_B = 3$. There are $N = 10000$ students and the school has capacity 0.1. To find the equilibrium $\tilde{q}^*_{\text{FULL}}$, we search over a grid of 250 threshold values. The mean over 20 simulation runs is presented, along with 95% confidence intervals.

Parameters for Figure EC.5. Figure EC.5 illustrates Proposition EC.1, which characterizes admission outcomes in a setting in which students make strategic decisions on whether to take the test. The figures show how the admitted students' (a) academic merit, (b) diversity level, and (c) individual fairness gap depend on test informativeness σ^2_K for Group B, as a function of either the test cost c_B or the student skill level q.

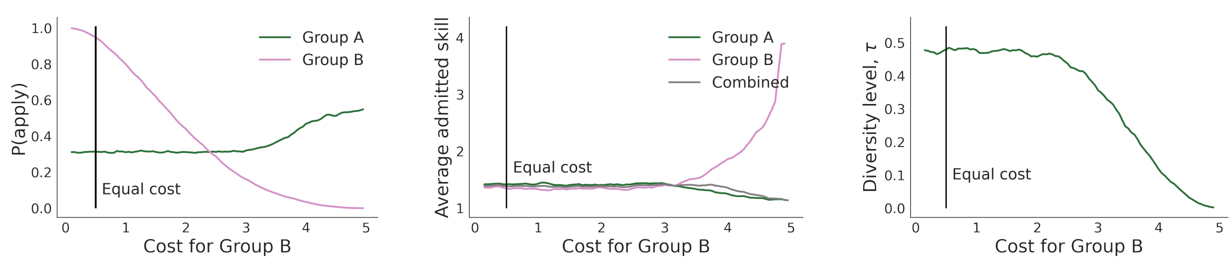
There are two features, where the where the non-test feature is equally informative for both groups, and compares instance a) where the test score is equally informative for both groups ($\sigma^2_{AK} = \sigma^2_{BK} = 1$) and b) where the test score is more informative for group A than B ($\sigma^2_{AK}1, \sigma^2_{BK} = 4$). Figures EC.5a and EC.5b fix cost $c_A = 0.5$ and vary $c_B \in [0, 5)$. Figure EC.5c considers $c_A = 0.5$ and a fixed cost $c_B = 3$. The remainder of the parameters are the same as Figure EC.4. The true skill distribution for both groups is Normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 1$. The features for the two groups are Normally distributed with mean $\mu_{gk} = 0$ for all g, k and $\sigma^2_{A1} = \sigma^2_{A2} = \sigma^2_{B1} = 1$. Students of both groups have valuation $v = 5$ for the school. Test costs are $c_A = 0.5$ and $c_B = 3$. There are $N = 10000$



(a) Academic merit (b) Diversity level (c) Individual fairness gap

Figure EC.5 Simulation results illustrating Proposition EC.1, characterizing admission outcomes in a setting in which students make strategic decisions on whether to take the test. The figures show how the admitted students' (a) academic merit, (b) diversity level, and (c) individual fairness gap depend on test informativeness σ^2_{BK} for Group B, as a function of either the test cost c_B or the student skill level q. Figures (a) and (b) fix cost $c_A = 0.5$ and vary c_B. Academic merit and diversity are particularly harmed when the test is costly and informative for Group B. Figure (c) considers a fixed cost $c_B = 3$ and shows that individual fairness is worse when the test is more informative for Group B. The full parameter set can be found in Appendix EC.2.2.1. Overall, when the test is informative (low feature variance), higher costs for group B correspond to worse outcomes across all metrics.

students and the school has capacity 0.1. To find the equilibrium $\tilde{q}^*_{\text{FULL}}$, we search over a grid of 250 threshold values. The mean over 20 simulation runs is presented, along with 95% confidence intervals.



(a) Probability of applying (b) Academic merit (c) Diversity level

Figure EC.6 Strategic students setting. How the probability of applying and the admitted students' academic merit change when Group A has cost $c_A = 0.5$ and the cost for Group B varies. As the cost for Group B increases, fewer Group B students apply and more Group A students apply (since the threshold decreases). Academic merit of admitted Group B students increases while that of Group A decreases. We consider a setting where the variances of non-test features are equal for both groups, but Group B has higher variance for test; the full parameter set can be found in Electronic Companion EC.2.2.1.

Dropping the test score. Figure EC.7 shows the change in the diversity level and average skill level of the admitted students, after dropping the test. In this scenario, since the variance of the non-test feature $\sigma^2_{A0} = \sigma^2_{B0} = 1$ are equal for both groups, a test-free policy will have a diversity level of $\tau = 0.5$.

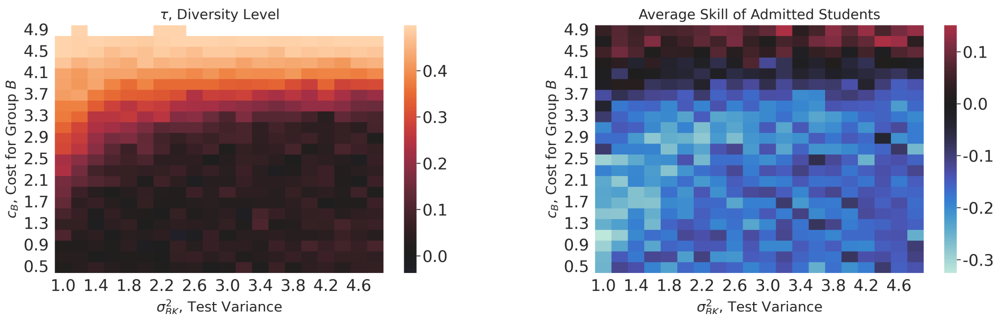


Figure EC.7 Change in diversity and average skill when the school drops the test requirement. When the test cost is large enough for Group B, dropping the test requirement increases the average academic merit and the diversity of the admitted student body. The full parameter set can be found in Electronic Companion EC.2.2.1.

EC.2.3. Simulations with synthetic data and two schools

The simulation setup closely resembles that of the single school, strategic student setting outlined in EC.2.2.1. In the same way as the single school, strategic student setting, each student is initialized with their true skill q and non-test features $\boldsymbol{\theta}_{\text{SUB}}$. The student observes $\boldsymbol{\theta}_{\text{SUB}}$ and makes their test decision by calculating their expected reward for the situation in which they take the test and the situation in which they do not.

The simulation setup with two schools J_1 and J_2 differs in that we now have two testing policies (P^1, P^2) and two admission thresholds $(\tilde{q}_1^*, \tilde{q}_2^*)$. For each policy pair –$(P^1_{\text{FULL}}, P^2_{\text{FULL}})$, $(P^1_{\text{FULL}}, P^2_{\text{SUB}})$, $(P^1_{\text{SUB}}, P^2_{\text{FULL}})$, and $(P^1_{\text{SUB}}, P^2_{\text{SUB}})$—we simulate the resulting admission outcomes in equilibrium.

Given a fixed policy pair (P^1, P^2) and admission threshold pair $(\tilde{q}_1^*, \tilde{q}_2^*)$, students observe their non-test features $\boldsymbol{\theta}_{\text{SUB}}$ and calculates the distribution of their estimated skill if they were to take the test $\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}$, in the same way that they do in the single school, strategic setting. Then, the student solves for their optimal test decision. For example, if the policy pair is $(P^1_{\text{FULL}}, P^2_{\text{SUB}})$, then the student solves the following optimization problem:

$$\alpha(\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g; \mathbf{P}) = \arg\max_{\alpha \in \{0,1\}} \alpha \left(v_1 \mathbb{P}(Y_1 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^1_{\text{FULL}}) - c_g\right) + v_2 \mathbb{P}(Y_1 = 0 \cap Y_2 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^2_{\text{SUB}}).$$

The student then applies to all schools that do not require the test and applies to a test-required school if they choose $\alpha = 1$.

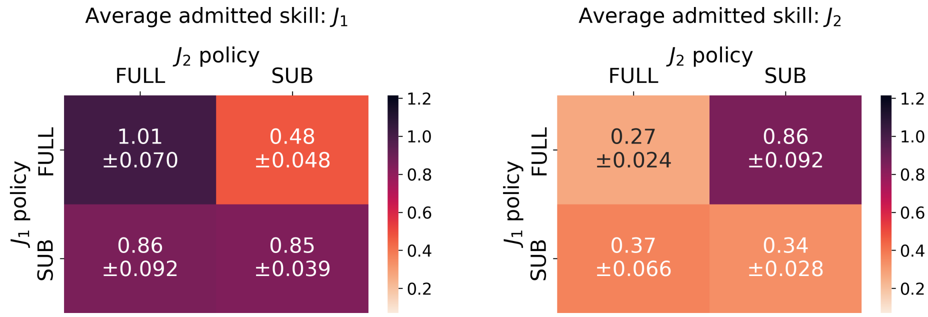


Figure EC.8 Average admitted skill of the resulting student body, in J_1 (preferred school) and J_2 (less preferred school), as a function of J_1 and J_2 testing policies. Students in both groups have valuations $v_1 = 3$ and $v_2 = 2$ and students have test cost $c_A = c_B = 1.5$. See EC.2.3.1 for full parameter details.

For each policy pair (P^1, P^2), finding an equilibrium amounts to finding an admission threshold pair $(\tilde{q}_1^*, \tilde{q}_2^*)$ such that each school admits the largest number of students while respecting their capacity constraints. We do a grid search to find the equilibrium admission threshold pair $(\tilde{q}_1^*, \tilde{q}_2^*)$.

EC.2.3.1. Parameters for the figures in the main text Figure 6 shows the average admitted skill (academic merit) of the resulting student body in J_1 (the preferred school) and J_2 (the less preferred school), as a function of J_1 and J_2 testing policies.

In both Figure 6a and Figure 6b, there are $N = 1000$ students, with half in group A and half in group B. The true skill is Normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 1$. Students have two features. Group A has feature distributions with mean $\mu_{A1} = \mu_{A2} = 1$ and variance $\sigma_{A1}^2 = \sigma_{A2}^2 = 1$. Group B has feature distributions with mean $\mu_{B1} = \mu_{B2} = -1$ and variance $\sigma_{B1}^2 = 3, \sigma_{B2}^2 = 5$.[17] Students have valuation $v_1 = 3$ for J_1 and $v_2 = 2$ for J_2. We simulate $N = 1000$ students. J_1 and J_2 each have capacity 0.2. In Figure 6a, the test costs are $c_A = c_B = 0.5$. In Figure 6b, the test costs are $c_A = c_B = 2.0$. To find the equilibrium threshold pair $(\tilde{q}_1^*, \tilde{q}_2^*)$ we do a grid search over 100 values of $\tilde{q}_1^*$ and 100 values of $\tilde{q}_2^*$.

Figure EC.4 shows simulation results illustrating the student equilibrium decisions (characterized by Equation (5)) on whether to take the test and apply to a school that requires the test, as a function of their true skill and group.

There are two features, where the where the non-test feature is equally informative for both groups, but the test score is more informative for group A than group B. The true

[17] Note that the skill estimates and admission outcomes depend only on the feature variance, not the mean, since the schools' Bayesian estimation method can account for the shift in the mean.

skill distribution for both groups is Normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 1$. The features for the two groups are Normally distributed with mean $\mu_{gk} = 0$ for all g, k and $\sigma^2_{A1} = \sigma^2_{A2} = \sigma^2_{B1} = 1$ and $\sigma^2_{B2} = 2$, where $k = 2$ denotes the test feature. Students of both groups have valuation $v = 5$ for the school. Test costs are $c_A = 0.5$ and $c_B = 3$. There are $N = 10000$ students and the school has capacity 0.1. To find the equilibrium, we search over a grid of 250 threshold values.

EC.2.3.2. Additional simulations for synthetic data and two schools. Figure EC.8 shows additional simulations for the two school, strategic student setting, under different utility and test costs parameters. In this setting, the optimal policies of both schools can depend on the policy of their competitor. J_1's optimal policy is to require the test when J_2 requires the test, but drop the test when J_2 drops the test. J_2's optimal policy is to drop the test when J_1 requires the test, but when J_1 drops the test, J_2 receives a quite similar average admitted skill when dropping or requiring the test.

EC.3. Supplementary information and discussion

Here, we provide additional information to support the main text analysis and writing. Table EC.3 provides a table of key notation. We also include further discussion on various modeling points.

Strategic student behavior in practice. Our model of rational student behavior requires that students know school cutoffs in equilibrium. In practice, there is substantial uncertainty about school admission policies across application settings (Tomkins et al. 2023, Idoux 2023, Kapor et al. 2020, Ajayi and Sidibe 2020), and students may behave suboptimally given knowledge of historical college decisions (Tomkins et al. 2023). It may be possible to incorporate such behavioral and informational effects into the model, using ideas from application search models under imperfect information (Ajayi and Sidibe 2020, Calsamiglia et al. 2020, Agarwal and Somaini 2018, 2020, Idoux 2023). Our results provide qualitative, directional insight regarding student behavioral effects. For example, we expect the results to continue to hold in settings where student beliefs regarding their admission chances is monotonic increasing in their test scores and their knowledge of their other features; however, showing such a result would require moving beyond our distributional assumptions and specifying a specific search model or belief structure for students.

Table EC.3 Key mathematical notation

Symbol	Meaning	Section
q	Student's latent (unobserved) skill level	Section 2 (Base Model)
θ_k	Feature k (e.g., test score, grades, etc.)	Section 2 (Base Model)
$\theta = (\theta_1, ..., \theta_K)$	Vector of all features	Section 2 (Base Model)
ϵ_k	Gaussian noise term for feature k	Section 2 (Base Model)
$g \in \{A, B\}$	Student group (A or B)	Section 2 (Base Model)
π	Mass/proportion of students in group B	Section 2 (Base Model)
μ	Mean of skill distribution	Section 2 (Base Model)
σ^2	Variance of skill distribution	Section 2 (Base Model)
μ_{gk}	Mean of noise distribution for feature k and group g	Section 2 (Base Model)
σ^2_{gk}	Variance of noise distribution for feature k and group g	Section 2 (Base Model)
γ_g	Fraction of group g with access to full set of features	Section 2 (Base Model)
$\tilde{q}(\theta, g)$	Perceived skill estimate given features θ and group g	Section 3.1 (Bayesian Estimation)
$\tilde{q}^*_S$	Admission threshold under policy P_S	Section 3.1 (Bayesian Estimation)
$Y \in \{0, 1\}$	Admission decision (1 = admitted)	Section 2 (Base Model)
$\tau(P)$	Diversity level (fraction of admitted students from group B) under policy P	Section 4 (Analysis)
$I(q; P)$	Individual fairness gap at skill level q under policy P	Section 4 (Analysis)
c_g	Cost for group g to take test (in strategic model)	Section 5 (Strategic Model)
v	Value/utility of admission (in strategic model)	Section 5 (Strategic Model)
$\alpha \in \{0, 1\}$	Student's action (1 = apply/take test)	Section 5 (Strategic Model)
P_{FULL}	Policy requiring full set of features	Section 4.2 (Policy Analysis)
P_{SUB}	Policy requiring only subset of features	Section 4.2 (Policy Analysis)
$\underline{q}^g$	Threshold for taking test in strategic model	Section 5.1 (Single School)

Model with general competition across many schools and arbitrary student preferences. Our multi-school analysis is restricted to studying two schools, where all applicants prefer one program over another. Studying competition more generally would be of interest, such as when student preferences are heterogeneous (students differ in which schools they relatively prefer). We note that our results suggest that the homogeneous setting already induces competition: the best response policy of the more-preferred school can depend on the policy of the less-preferred school; intuitively, students may not choose to take the test if they can be admitted to the less-preferred school without the test. The policies of the schools jointly affect student strategic incentives, and in turn school optimal policies. We foresee that analyzing more general competition effects would use substantially different technical tools (and likely start from a different model than our base model).

EC.4. Proofs of statements

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

EC.4.1. Auxiliary lemmas

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

Lemma EC.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 EC.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 EC.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 EC.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 EC.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 EC.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:

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

which is negative for any $y \in \mathbb{R}$. Therefore, $\frac{\mathrm{d}\hat{h}(y)}{\mathrm{d}y} < 0$ 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 $x > 0$. □

EC.4.2. Group-aware estimation (Proofs from Section 3)

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

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

Perceived skill conditional on true skill. Equation (EC.1) gives us the skill estimate $\tilde{q}$ of a student conditional on features $\boldsymbol{\theta}$. Another useful distribution is $\tilde{q} \mid q, g, P_S$, which is also Gaussian. Indeed, observe that $\tilde{q}(\boldsymbol{\theta}, g)$ in Equation (EC.1) is a linear combination of independent (conditional on q) Gaussian variables $\theta_k = q + \epsilon_{kg}$, $k \in S$. Thus,

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

Lemma EC.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. □

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 EC.7 **(Lemma 1)**. *Consider a school that uses feature set $S \subseteq \{1,\dots,K\}$ 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 EC.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:

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

□

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

Corollary EC.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 EC.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*

$$q \mid \tilde{q}, g, P_S \sim \mathcal{N}\left(\tilde{q}, \frac{1}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right). \tag{EC.3}$$

Proof. We apply Lemma EC.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 EC.2, we get that

$$\begin{aligned}
\mathbb{E}[M \mid \tilde{q}, g, P_S] &= \frac{\sigma^2 \frac{(\sum_{k\in S}\sigma_{gk}^{-2})^2}{(\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2})^2}\tilde{q} + \mu\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{(\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2})^2}}{\sigma^2 \frac{(\sum_{k\in S}\sigma_{gk}^{-2})^2}{(\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2})^2} + \frac{\sum_{k\in S}\sigma_{gk}^{-2}}{(\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2})^2}} \\
&= \frac{\sum_{k\in S}\sigma_{gk}^{-2}\tilde{q} + \mu\sigma^{-2}}{\sum_{k\in S}\sigma_{gk}^{-2} + \sigma^{-2}} \\
\mathrm{Var}[M \mid \tilde{q}, g, P_S] &= \left(\left(\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2})^2}\right)^{-1} + \sigma^{-2}\left(\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}}\right)^{-2}\right)^{-1} \\
&= \frac{\left(\sum_{k\in S}\sigma_{gk}^{-2}\right)^2}{\left(\sigma^{-2} + \sum_{k\in S}\sigma_{gk}^{-2}\right)^3}.
\end{aligned}$$

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

EC.4.3. Baseline policy in the absence of barriers (Proofs from Sections 4.1)

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

$$(1-\pi)F_{\tilde{q}|A,P_S}(\tilde{q}_S^*)+\pi F_{\tilde{q}|B,P_S}(\tilde{q}_S^*) = 1-C. \tag{EC.4}$$

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, Equation (EC.4) can be equivalently written as

$$\Phi\left((\tilde{q}_S^*-\mu)\left((1-\pi)\sigma^2\left[\frac{\sum_{k\in S}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in S}\sigma_{Ak}^{-2}}\right]+\pi\sigma^2\left[\frac{\sum_{k\in S}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in S}\sigma_{Bk}^{-2}}\right]\right)^{-1/2}\right) = 1-C. \tag{EC.5}$$

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 $\sum_{k\in S}\sigma_{Ak}^{-2}$ 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 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, even in the absence of barriers. Given unequal precisions:*

(i) Diversity level*: Group B students are under-represented, i.e., $\tau(P_S)<\pi$. Furthermore, a 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., $I(q;P_S) > 0$, iff $q > \tilde{q}_S^* + \frac{\sigma^{-2}(\tilde{q}_S^* - \mu)}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}}$.*

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 $q > \mu + \sigma\Phi^{-1}(1-C)$, where Φ denotes the CDF of $\mathcal{N}(0,1)$.

(iii) Academic merit*: Admitted group B students have lower academic merit than group A.*

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 $F_{\tilde{q}|A,P_S}(\tilde{q}_S^*) = F_{\tilde{q}|B,P_S}(\tilde{q}_S^*) = 1-C$. 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 EC.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 Equation (EC.4) 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 $0 \le \Delta < \Delta'$. 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 Equation (EC.5). Solving for $\tilde{q}_S^*(\Delta)$ gives us

$$\tilde{q}_S^*(\Delta) = \mu + \Phi^{-1}(1-C)\cdot\left((1-\pi)\sigma^2\left[\frac{\sum_{k\in S}\sigma_{Bk}^{-2} + \Delta}{\sigma^{-2} + \sum_{k\in S}\sigma_{Bk}^{-2} + \Delta}\right] + \pi\sigma^2\left[\frac{\sum_{k\in S}\sigma_{Bk}^{-2}}{\sigma^{-2} + \sum_{k\in S}\sigma_{Bk}^{-2}}\right]\right)^{1/2}, \tag{EC.6}$$

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

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

□

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

$$\mathbb{P}[Y=1\mid q,g,P_S]=1-F_{\tilde{q}|q,g,P_S}(\tilde{q}_S^*)=1-\Phi\left(\frac{\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{gk}^{-2}}}\left(\tilde{q}_S^*-\frac{\mu\sigma^{-2}+q\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2}}\right)\right), \tag{EC.7}$$

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

$$\begin{aligned}
&\frac{\sigma^{-2}+\sum_{k\in S}\sigma_{Ak}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}}\left(\tilde{q}_S^*-\frac{\mu\sigma^{-2}+q\sum_{k\in S}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in S}\sigma_{Ak}^{-2}}\right)<\frac{\sigma^{-2}+\sum_{k\in S}\sigma_{Bk}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}}\left(\tilde{q}_S^*-\frac{\mu\sigma^{-2}+q\sum_{k\in S}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in S}\sigma_{Bk}^{-2}}\right)\\
\iff&\frac{\tilde{q}_S^*\sigma^{-2}+\tilde{q}_S^*\sum_{k\in S}\sigma_{Ak}^{-2}-\mu\sigma^{-2}-q\sum_{k\in S}\sigma_{Ak}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}}<\frac{\tilde{q}_S^*\sigma^{-2}+\tilde{q}_S^*\sum_{k\in S}\sigma_{Bk}^{-2}-\mu\sigma^{-2}-q\sum_{k\in S}\sigma_{Bk}^{-2}}{\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}}\\
\iff&\left(\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}-\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}\right)\left(\sigma^{-2}(\tilde{q}_S^*-\mu)+(q-\tilde{q}_S^*)\sqrt{\sum_{k\in S}\sigma_{Bk}^{-2}}\sqrt{\sum_{k\in S}\sigma_{Ak}^{-2}}\right)<0.
\end{aligned} \tag{EC.8}$$

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 Equation (EC.8) 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 Equation (EC.2), it follows that

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

By Equation (EC.6), 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 $\mathrm{Var}[\tilde q\mid q,A,P_S]$ 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. $\square$

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 EC.3, we get that

$$\begin{aligned}\mathbb{E}[\tilde q\mid Y=1,g,P_S]=\mathbb{E}[\tilde q\mid \tilde q\geq\tilde q_S^*,g,P_S]&=\mathbb{E}[\tilde q\mid g]+\sqrt{\mathrm{Var}[\tilde q\mid g,P_S]}\frac{\phi(t_g)}{1-\Phi(t_g)}\\&=\mu+\sigma\sqrt{\frac{\sum_{k\in S}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in S}\sigma_{gk}^{-2}}}\cdot\frac{\phi(t_g)}{1-\Phi(t_g)},\end{aligned}\tag{EC.9}$$

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 EC.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 EC.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,

$$\mathbb{E}[q \mid Y=1, g, P_S] = \mathbb{E}[q \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S] = \mathbb{E}_{\tilde{q}}[\mathbb{E}_q[q \mid \tilde{q}, g, P_S] \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S] = \mathbb{E}[\tilde{q} \mid \tilde{q} \geq \tilde{q}_K^*, g, P_S], \tag{EC.10}$$

where the last equality follows from Lemma EC.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 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 $\mathbb{E}[q \mid Y=1, A, P_S] > \mathbb{E}[q \mid Y=1, B, P_S]$. □

Explaining why the individual fairness gap decreases for high-skilled students. 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 EC.9b.

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 Equation (EC.2), 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

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

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

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

The above condition is equivalent to

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

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 $\lim_{q\to\infty}I(q;P_S)=0$. □

EC.4.4. Dropping a feature with and without barriers (Proofs from Section 4.2)

Dropping a feature in the absence of barriers. We are interested in comparing the 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 EC.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. □

Let $\tilde{q}^*_{\text{SUB}}$ be the decision threshold of a school considering only features $k = 1$ to $K-1$. By Equation (EC.4), $\tilde{q}^*_{\text{SUB}}$ is the solution to the following equation

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

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

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

Lemma EC.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 EC.9. □

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 the test, $\tau(P_{\text{SUB}}) > \tau(P_{\text{FULL}})$, iff*

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

(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} iff $q > q_g$. 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 $g \in \{A, B\}$, that is, $\mathbb{E}[q \mid Y = 1, g, P_{\text{FULL}}] > \mathbb{E}[q \mid Y = 1, g, P_{\text{SUB}}]$.*

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 Equation (EC.5), 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 Equation (3). $\square$

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 Equation (EC.5) in Equation (EC.7), 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

$$\begin{aligned}(\mu - q)&\left(\sqrt{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}} - \sqrt{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}\right) + \sigma\Phi^{-1}(1-C)\sqrt{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}\sqrt{(1-\pi)\frac{\frac{\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{Bk}^{-2}}} + \pi} \\ &- \sigma\Phi^{-1}(1-C)\sqrt{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}\sqrt{(1-\pi)\frac{\frac{\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Ak}^{-2}}}{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}} + \pi} > 0.\end{aligned} \tag{EC.11}$$

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

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

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 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

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

Let

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

Define $\hat{q} \triangleq \max\{\underline{q}, \underline{\underline{q}}, q_A, q_B\}$. 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, $I(q; P_{\text{FULL}}) > I(q; P_{\text{SUB}})$. $\square$

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 EC.10, $\tilde{q}^*_{\text{SUB}} < \tilde{q}^*_{\text{FULL}}$, the expected estimated skill of each admitted group decreases, that is,

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

Equation (EC.10) further implies that $\mathbb{E}[q \mid Y=1, g, P_{\text{SUB}}] < \mathbb{E}[q \mid Y=1, g, P_{\text{FULL}}]$. $\square$

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

$$(1-\pi)\gamma_A(1 - F_{\tilde{q}|A,P_{\text{FULL}}}(\tilde{w}^*_{\text{FULL}})) + \pi\gamma_B(1 - F_{\tilde{q}|B,P_{\text{FULL}}}(\tilde{w}^*_{\text{FULL}})) = C. \tag{EC.12}$$

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

THEOREM EC.1 (**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 $\Delta_g(\xi_g, \rho^g_{\text{SUB}})$ such that the academic merit of group g increases if and only if

$$\beta_g(\gamma_A, \gamma_B, \rho^g_{\text{FULL}}) \leq \Delta_g(\xi_g, \rho^g_{\text{SUB}}), \tag{EC.13}$$

where

$$\rho^A_S = \frac{1}{\rho^B_S} \triangleq \frac{\sum_{k\in S}\sigma^{-2}_{Bk}}{\sigma^{-2} + \sum_{k\in S}\sigma^{-2}_{Bk}}\left(\frac{\sum_{k\in S}\sigma^{-2}_{Ak}}{\sigma^{-2} + \sum_{k\in S}\sigma^{-2}_{Ak}}\right)^{-1},$$

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

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

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

As barriers to group g increase (γ_g decreases), then $\beta_g(\gamma_A, \gamma_B, \rho^g_{\text{SUB}})$ 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 $\eta(1,1,\rho^B_{\text{SUB}}) > \eta(\gamma_A, \gamma_B, \rho^B_{\text{FULL}})$, where

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

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 Equation (EC.13) *holds.* We adopt an argument similar to the proof of Proposition EC.7. We prove the statement for $g = A$. The argument for group B is similar.

First, similarly to Equation (EC.5), we derive that

$$\tilde{w}^*_{\text{FULL}} = \mu + \Phi^{-1}\left(1 - \frac{C}{(1-\pi)\gamma_A + \pi\gamma_B}\right)\sigma\sqrt{\frac{(1-\pi)\gamma_A \frac{\sum_{k\in\text{FULL}}\sigma^{-2}_{Ak}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma^{-2}_{Ak}} + \pi\gamma_B\frac{\sum_{k\in\text{FULL}}\sigma^{-2}_{Bk}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma^{-2}_{Bk}}}{(1-\pi)\gamma_A + \pi\gamma_B}}. \tag{EC.14}$$

Second, requiring that $\mathbb{E}[q \mid Y = 1, A, P_{\text{FULL}}] \le \mathbb{E}[q \mid Y = 1, A, P_{\text{SUB}}]$ and adapting Lemma EC.3 to our setting with barriers gives us

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

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

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

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 $\beta_A(\gamma_A, \gamma_B, \rho^A_{\text{FULL}}) \leq \Delta_A(\xi_A, \rho^A_{\text{SUB}})$.

Step 2: We show that, for each group $g \in \{A, B\}$, $\beta_g(\gamma_g, \gamma_{g'}, \rho^g_{\text{FULL}})$ 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:

$$\sqrt{\frac{\frac{\pi_{g'}\gamma_{g'}}{\pi_g\gamma_g}\rho^g_S + 1}{1 + \frac{\pi_{g'}\gamma_{g'}}{\pi_g\gamma_g}}}. \tag{EC.15}$$

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 Equation (EC.15) with respect to γ_A equals

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

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

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 $\beta_A(\bar{\bar{\gamma}}_A, \gamma_B, \rho^A_{\text{FULL}}) = \Delta_A(\xi_A, \rho^A_{\text{SUB}})$ 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 $\beta_A(1, \gamma_B, \rho^A_{\text{FULL}}) > \Delta_A(\xi_A, \rho^A_{\text{FULL}})$. Furthermore, observe that for $\gamma_A = 0$, academic merit trivially improves from $\beta_A(0, \gamma_B, \rho^A_{\text{FULL}}) = 0$ to a positive value $\Delta_A(\xi_A, \rho^A_{\text{SUB}}) > 0$ after dropping test scores. Thus, by the continuity of $\beta_A(\gamma_A, \gamma_B, \rho^A_{\text{FULL}})$, 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 $\beta_A(\gamma_A, \gamma_B, \rho^B_{\text{FULL}}) < \Delta_A(\xi_A, \rho^A_{\text{SUB}})$ for all $\gamma_A \in [0, \epsilon)$. Consequently, $\bar{\bar{\gamma}}_A \geq \epsilon > 0$. $\square$

Proof of Part (ii). Plugging Equation (EC.14) into the definition of diversity with and without test scores, respectively, it immediately follows that diversity improves if and only if $\eta(1, 1, \rho^B_{\text{SUB}}) > \eta(\gamma_A, \gamma_B, \rho^B_{\text{FULL}})$.

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., $\eta(\gamma_A, \gamma'_B, \rho^B_{\text{FULL}}) > \eta(\gamma_A, \gamma_B, \rho^B_{\text{FULL}})$ for $\gamma'_B > \gamma_B$.

By Equation (EC.14), 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 $\eta(1,1,\rho^B_{\text{SUB}}) = \eta(\gamma_A, \bar{\gamma}, \rho^B_{\text{FULL}})$ and (b) $\bar{\gamma} \in (0,1]$. The proof follows as in Step 3 in Part (i). $\square$

EC.4.5. Strategic students: Single school (Proofs from Section 5.2.1)

LEMMA EC.11. *Fix testing policy P_{FULL}. Let $\alpha(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) : \mathbb{R} \times \{A,B\} \to \{0,1\}$ denote the function that describes the action of students in group g with skill estimate $\tilde{q}_{\text{SUB}}$, i.e.,*

$$\alpha(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) \triangleq \arg\max_{\alpha \in \{0,1\}} \alpha \left(v\, \mathbb{P}(Y=1 \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}) - c_g\right). \tag{EC.16}$$

At equilibrium, for any $\boldsymbol{\theta}_{\text{SUB}} \in \mathbb{R}^{K-1}$ and $g \in \{A,B\}$, it holds that

$$\alpha(\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g; P_{\text{FULL}}) = \arg\max_{\alpha \in \{0,1\}} \alpha \left(v\, \mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) - c_g\right)$$

Proof. Recall that $\tilde{q}^*_{\text{FULL}}$ denotes the admission threshold of the school with policy P_{FULL} at a given equilibrium. To solve Equation (5), the student computes the following probability:

$$\begin{aligned}
\mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) &= \mathbb{P}(\tilde{q}(\boldsymbol{\theta}_{\text{FULL}}, g) \geq \tilde{q}^*_{\text{FULL}} \mid \boldsymbol{\theta}_{\text{SUB}}, g) \\
&= \mathbb{P}_{\theta_K}\left(\frac{\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g)(\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}) + (\theta_K - \mu_{gK})\sigma_{gK}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}} \sigma_{gk}^{-2}} \geq \tilde{q}^*_{\text{FULL}} \mid \boldsymbol{\theta}_{\text{SUB}}, g \right) \\
&= \mathbb{P}_{\theta_K}\left(\theta_K \geq \mu_{gK} + \tilde{q}^*_{\text{FULL}} + \sigma_{gK}^2(\tilde{q}^*_{\text{FULL}} - \tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g))(\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}) \mid \boldsymbol{\theta}_{\text{SUB}}, g \right) \\
&= \mathbb{P}_{\theta_K}\left(\theta_K \geq \mu_{gK} + \tilde{q}^*_{\text{FULL}} + \sigma_{gK}^2(\tilde{q}^*_{\text{FULL}} - \tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g))(\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}) \mid \tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g \right) \\
&= \mathbb{P}(Y=1 \mid \tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g, P_{\text{FULL}}),
\end{aligned}$$

where in the second line we used Equation (EC.1) for $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{FULL}}, \boldsymbol{\theta}_{\text{SUB}}$ to rewrite $\tilde{q}(\boldsymbol{\theta}_{\text{FULL}}, g)$ in terms of $\tilde{q}_{\text{SUB}}(\boldsymbol{\theta}_{\text{SUB}})$ and θ_K, i.e.,

$$\begin{aligned} \tilde{q}(\boldsymbol{\theta}_{\text{FULL}}, g) &= \frac{\mu\sigma^{-2} + \sum_{k\in\text{FULL}}(\theta_k - \mu_{gk})\sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}}\sigma_{gk}^{-2}} \\ &= \frac{\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g)(\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{gk}^{-2}) + (\theta_K - \mu_{gK})\sigma_{gK}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}. \end{aligned} \tag{EC.17}$$

This equality immediately implies that for any $\alpha \in \{0,1\}$:

$$\alpha(v\,\mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) - c_g) = \alpha(v\,\mathbb{P}(Y=1 \mid \tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g, P_{\text{FULL}}) - c_g).$$

Consequently,

$$\begin{aligned} \arg\max_{\alpha\in\{0,1\}} \alpha(v\,\mathbb{P}(Y=1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P_{\text{FULL}}) - c_g) &= \arg\max_{\alpha\in\{0,1\}} \alpha(v\,\mathbb{P}(Y=1 \mid \tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g, P_{\text{FULL}}) - c_g) \\ &= \alpha(\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g), g; P_{\text{FULL}}), \end{aligned}$$

which concludes the proof of the lemma. $\square$

LEMMA 3. *Suppose that the school uses a test-based policy P_{FULL}. There exists a unique equilibrium (α^*, Y^*), with the following property: there is a threshold $\underline{q}^g$ such that students in group g take the test $(a=1)$ if and only if $\tilde{q}_{\text{SUB}} \geq \underline{q}^g$, where*

$$\underline{q}^g = \tilde{q}^*_{\text{FULL}} - \Phi^{-1}\left(1 - \frac{c_g}{v}\right)\left(\frac{\sigma_{gK}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}\right)\sqrt{\sigma_{gK}^2 + \frac{1}{\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}, \tag{7}$$

and $\tilde{q}^*_{\text{FULL}}$ is the solution to Equation (6) so that $Y^*(\tilde{q}_{\text{FULL}}; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{\text{FULL}}\}$.

Proof. Without loss of generality, we fix group g throughout the proof as the arguments are analogous for both groups of students.

Step 1: We derive the distribution of $\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}$. The student uses this distribution to solve Equation (EC.16).

Fix test-free skill estimate $\tilde{q}_{\text{SUB}}$. By Lemma EC.8, we have that

$$q \mid \tilde{q}_{\text{SUB}}, g \sim \mathcal{N}\left(\tilde{q}_{\text{SUB}}, \frac{1}{\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}\right). \tag{EC.18}$$

Furthermore, conditional on her true skill q, the student's test score θ_K is drawn from a distribution $\theta_K \mid q, g \sim \mathcal{N}(q + \mu_{gK}, \sigma_{gK}^2)$. Applying Lemma EC.1, we get that

$$\theta_K \mid \tilde{q}_{\text{SUB}}, g \sim \mathcal{N}\left(\tilde{q}_{\text{SUB}} + \mu_{gK}, \sigma_{gK}^2 + \frac{1}{\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}\right). \tag{EC.19}$$

By applying Lemma EC.1, combined with Equation (EC.17) and the above distribution, the student then finds that her projected skill estimate $\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}$, after they take the test and submit the score θ_K to the school, will follow a Normal distribution:

$$\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}} \sim \mathcal{N}\left(\tilde{q}_{\text{SUB}}, \left(\frac{\sigma_{gK}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}\right)^2 \left(\sigma_{gK}^2 + \frac{1}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}\right)\right). \tag{EC.20}$$

Step 2: Neither $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = 1$, $\forall\ \tilde{q}_{\text{SUB}}$, or $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = 0$, $\forall\ \tilde{q}_{\text{SUB}}$, constitute an equilibrium. For the sake of contradiction, assume that $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = 1$, $\forall\ \tilde{q}_{\text{SUB}}$, is an equilibrium. Then, all students take the test and apply to the school as in the main setup without barriers.

The student has probability $\mathbb{P}(Y = 1 \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}) = \mathbb{P}(\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}})$ to be accepted by the school. Keeping $\tilde{q}^*_{\text{FULL}}$ fixed, by Equation (EC.20), there exists a small enough $\underline{q}$ such that for all $\tilde{q}_{\text{SUB}} < \underline{q}$, $v\,\mathbb{P}(Y = 1 \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}}) - c_g < 0$. Thus, students with $\tilde{q}_{\text{SUB}} < \underline{q}$ have incentive not to apply, implying that $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = 0$ for a positive mass of students, which contradicts our assumption.

A similar argument also shows that $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = 0$ cannot be an equilibrium, since students with $\tilde{q}_{\text{SUB}} > \overline{q}$ for some threshold $\overline{q}$ will have the incentive to deviate and take the test.

Step 3: If $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}})$ is an equilibrium student strategy, then it must be non-decreasing in $\tilde{q}_{\text{SUB}}$. We prove this claim by contradiction. Suppose that there exist $\tilde{q}'_{\text{SUB}}, \tilde{q}''_{\text{SUB}}$, with $\tilde{q}'_{\text{SUB}} < \tilde{q}''_{\text{SUB}}$, such that $1 = \alpha^*(\tilde{q}'_{\text{SUB}}, g; P_{\text{FULL}}) > \alpha^*(\tilde{q}''_{\text{SUB}}, g; P_{\text{FULL}}) = 0$.

We show that this cannot hold true. Indeed, since the mean of Equation (EC.20) is increasing in $\tilde{q}_{\text{SUB}}$ and the variance does not depend on $\tilde{q}_{\text{SUB}}$, it follows that

$$\mathbb{P}(Y = 1 \mid \tilde{q}'_{\text{SUB}}, g, P_{\text{FULL}}) \leq \mathbb{P}(Y = 1 \mid \tilde{q}''_{\text{SUB}}, g, P_{\text{FULL}}),$$

therefore $0 \leq v\,\mathbb{P}(Y = 1 \mid \tilde{q}'_{\text{SUB}}, g, P_{\text{FULL}}) - c_g \leq v\,\mathbb{P}(Y = 1 \mid \tilde{q}''_{\text{SUB}}, g, P_{\text{FULL}}) - c_g$, where the first inequality follows from the fact that $\alpha^*(\tilde{q}'_{\text{SUB}}, g; P_{\text{FULL}}) = 1$. Consequently, the student with $\tilde{q}''_{\text{SUB}}$ also has the incentive to apply, i.e., $\alpha^*(\tilde{q}''_{\text{SUB}}, g; P_{\text{FULL}}) = 1$ which is a contradiction. Thus, $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}})$ must be non-decreasing in $\tilde{q}_{\text{SUB}}$.

Step 4: If an equilibrium exists, it takes a threshold form: $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{SUB}} \geq \underline{q}^g\}$. An immediate corollary of Steps 2 and 3 is that if an equilibrium $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}})$ exists, it must take a threshold form, i.e., there must exist a threshold $\underline{q}^g$ such that $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{SUB}} \geq \underline{q}^g\}$. In other words, $\underline{q}^g$ corresponds to the unique skill level that characterizes students who are indifferent between taking and not taking the test.

Step 5: An equilibrium (α^*, Y^*) *exists and is unique.* As explained in the main text, the selection policy Y^* of the school remains the same as in the baseline setting without test costs: among the students who apply, the school sets a threshold $\tilde{q}^*_{\text{FULL}}$ to accept the top mass C of applicants thus $Y^*(\tilde{q}^*_{\text{FULL}}; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{\text{FULL}}\}$ where $\tilde{q}^*_{\text{FULL}}$ is the unique solution to:

$$\tilde{q}^*_S = \min\left\{z \in \mathbb{R} : \sum_g \pi_g \, \mathbb{E}_{\boldsymbol{\theta}_S}[\alpha^*(\boldsymbol{\theta}_{\text{SUB}}, g; P_S) \mid \tilde{q}(\boldsymbol{\theta}_S, g) \geq z, g, P_S] \leq C\right\}. \tag{EC.21}$$

Regarding α^*, we will prove the slightly more general statement: given any threshold $\tilde{q}^*_{\text{FULL}}$, there exists a unique equilibrium with $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{SUB}} \geq \underline{q}^g\}$ where $\underline{q}^g$ is the solution to

$$\mathbb{P}(\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{\text{FULL}} \mid \underline{q}^g, \boldsymbol{\theta}_{\text{SUB}}, g) = \frac{c_g}{v}. \tag{EC.22}$$

Indeed, given the admission cutoff $\tilde{q}^*_{\text{FULL}}$ and using Equation (EC.20), the student computes her admission probability:

$$\mathbb{P}(\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{\text{FULL}} \mid \underline{q}^g, g, P_{\text{FULL}}) = 1 - \Phi\left(\frac{\tilde{q}^*_{\text{FULL}} - \underline{q}^g}{\text{Var}(\tilde{q}_{\text{FULL}} \mid \underline{q}^g, g, P_{\text{FULL}})}\right).$$

Given that the CDF Φ is a continuous, strictly increasing function in $\tilde{q}_{\text{SUB}}$ and $c_g/v < 1$, it follows that Equation (EC.22) has a unique solution $\underline{q}^g$. Then, $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{SUB}} \geq \underline{q}^g\}$ is an equilibrium: all students with $\tilde{q}_{\text{SUB}} \geq \underline{q}^g$ receive weakly positive expected utility if they apply, whereas all students with $\tilde{q}_{\text{SUB}} < \underline{q}^g$ get strictly negative expected utility therefore they choose not to apply. By the uniqueness of the solution $\underline{q}^g$ to Equation (EC.22), it follows that no other equilibrium of a threshold form can exist. Due to Step 4, this further implies that $\alpha^*(\tilde{q}_{\text{SUB}}, g; P_{\text{FULL}}) = \mathbf{1}\{\tilde{q}_{\text{SUB}} \geq \underline{q}^g\}$ must be unique. Extending the arguments to students of any g concludes the proof. $\square$

EC.4.5.1. Effect of test cost and informativeness on admissions In the non-strategic setting of Proposition 1, the sign of the informativeness gap $\sum_{k\in S} \sigma_{Ak}^{-2} - \sum_{k\in S} \sigma_{Bk}^{-2}$ determined the diversity level and academic merit in a straightforward manner: if group B has lower total precision than group A, then it is under-represented and has lower academic merit among the admitted class. The same holds even in the presence of barriers as long as $\gamma_A \geq \gamma_B$. With costly testing, however, Proposition EC.1 below shows that the relationship between informativeness and fairness becomes more complex, depending on the costs and informativeness of the features with and without the test score. Recall that $\Phi_2(x, y; \rho)$ denotes the CDF of the standard bivariate Normal distribution with correlation ρ.

PROPOSITION EC.1. *Consider the equilibrium under policy P_{FULL}.*

(i) Diversity level*: Group B students are under-represented, i.e., $\tau(P_{\text{FULL}}) < \pi$, if and only if*

$$\frac{\Phi\left(\frac{a_A+b_A\mu}{\sqrt{1+\tilde{\sigma}_A^2 b_A^2}}\right) - \Phi_2\left(\frac{a_A+b_A\mu}{\sqrt{1+\tilde{\sigma}_A^2 b_A^2}}, \frac{\tilde{q}^*_{\text{FULL}}-\mu}{\tilde{\sigma}_A}; -\frac{\tilde{\sigma}_A b_A}{\sqrt{1+\tilde{\sigma}_A^2 b_A^2}}\right)}{\Phi\left(\frac{a_B+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) - \Phi_2\left(\frac{a_B+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{\text{FULL}}-\mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)} > \frac{\tilde{\sigma}_B}{\tilde{\sigma}_A}, \tag{EC.23}$$

where

$$\tilde{\sigma}_g = \sigma\sqrt{\frac{\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}},$$

$$a_g \triangleq a_g(\tilde{q}^*_{\text{FULL}}) = \frac{\frac{\sigma_{gK}^{-2}\Phi^{-1}\left(1-\frac{c_g}{v}\right)}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}\sqrt{\sigma_{gK}^2+\frac{1}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}+\mu\frac{\sigma^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}-\tilde{q}^*_{\text{FULL}}}{\frac{\sqrt{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}\sqrt{1+\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}}},$$

$$b_g = \sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}(\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2})}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}}.$$

(ii) Academic merit*: Policy P_{FULL} achieves worse academic merit for group B than group A if and only if $\lambda(a_A, b_A, \tilde{\sigma}_A, \tau_A) > \lambda(a_B, b_B, \tilde{\sigma}_B, \tau_B)$, where*

$$\lambda(a_g, b_g, \tilde{\sigma}_g, \tau_g) \triangleq \mu - \frac{\tilde{\sigma}_g^2 b_g}{\tau_g\sqrt{1+\tilde{\sigma}_g^2 b_g^2}}\phi\left(\frac{a_g+b_g\mu}{\sqrt{1+\tilde{\sigma}_g^2 b_g^2}}\right)\Phi\left(\tilde{q}^*_{\text{FULL}}\sqrt{1+\tilde{\sigma}_g^2 b_g^2}+\frac{b_g\tilde{\sigma}_g(a_g+b_g\mu)}{\sqrt{1+\tilde{\sigma}_g^2 b_g^2}}\right)$$
$$+\Phi(a_g+b_g\mu+b_g\tilde{\sigma}_g\tilde{q}^*_{\text{FULL}})\frac{\tilde{\sigma}_g^2\phi(\tilde{q}^*_{\text{FULL}})}{\tau_g}.$$

Proof of Part (i). Fix group g. We break the proof into steps.

Step 1: We derive the distribution of $\tilde{q}_{\text{SUB}} \mid \tilde{q}_{\text{FULL}}, g$. By Lemma EC.8, we have that

$$q \mid \tilde{q}_{\text{FULL}}, g \sim \mathcal{N}\left(\tilde{q}, \frac{1}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}\right),$$

while by Equation (EC.2),

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

Applying Lemma EC.1 gives us

$$\tilde{q}_{\text{SUB}} \mid \tilde{q}_{\text{FULL}}, g \sim \mathcal{N}\left(\frac{\mu\sigma^{-2} + \tilde{q}_{\text{FULL}} \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}, \frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{(\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2})^2}\left(1 + \frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}} \sigma_{gk}^{-2}}\right)\right).$$

Step 2: We show that

$$\tau_g = \frac{\pi_g \tilde{\sigma}_g}{C} \Phi\left(\frac{a_g + b_g \mu}{\sqrt{1 + \tilde{\sigma}_g^2 b_g^2}}\right) - \frac{\pi_g \tilde{\sigma}_g}{C} \Phi_2\left(\frac{a_g + b_g\mu}{\sqrt{1 + \tilde{\sigma}_g^2 b_g^2}}, \frac{\tilde{q}^*_{\text{FULL}} - \mu}{\tilde{\sigma}_g}; -\frac{\tilde{\sigma}_g b_g}{\sqrt{1 + \tilde{\sigma}_g^2 b_g^2}}\right),$$

where $\tilde{\sigma}_g, a_g, b_g$ are defined as above.

Given that the school's admission threshold is $\tilde{q}^*_{\text{FULL}}$, only students with $\tilde{q} \geq \tilde{q}^*_{\text{FULL}}$ get admitted. If no costs existed, the fraction of students who would get admitted under a fixed threshold $\tilde{q}^*_{\text{FULL}}$ would be

$$\int_{\tilde{q}^*_{\text{FULL}}}^{\infty} \phi\left(\frac{\tilde{q} - \mu}{\tilde{\sigma}_g}\right) \mathrm{d}\tilde{q},$$

by Lemma 1. However, in the presence of costs, by Lemma 3, among all students who in our continuum model could have $\tilde{q}_{\text{FULL}} > \tilde{q}^*_{\text{FULL}}$, only students with $\tilde{q}_{\text{SUB}} \geq \underline{q}^g_{\text{SUB}}$ apply. Conditional on having the same $\tilde{q}_{\text{FULL}}$, Step 1 implies that the fraction of applying students from group g equals

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

Consequently, putting everything together, we get that

$$\tau_g = \frac{\pi_g}{C} \int_{\tilde{q}^*_{\text{FULL}}}^{\infty} \Phi\left(\frac{\frac{\mu\sigma^{-2} + \tilde{q} \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}} - \underline{q}^g_{\text{SUB}}}{\frac{\sqrt{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}} \sqrt{1 + \frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}} \sigma_{gk}^{-2}}}}\right) \phi\left(\frac{\tilde{q} - \mu}{\tilde{\sigma}_g}\right) \mathrm{d}\tilde{q}. \tag{EC.24}$$

By Equation (10,010.1) in Owen (1980), we have that

$$\int_{-\infty}^{u} \Phi(a + bx) \phi\left(\frac{x - \mu}{\rho}\right) \mathrm{d}x = \rho \Phi_2\left(\frac{a + b\mu}{\sqrt{1 + \rho^2 b^2}}, \frac{u - \mu}{\rho}; -\frac{\rho b}{\sqrt{1 + \rho^2 b^2}}\right).$$

Furthermore, by Equation (10,010.1) in Owen (1980),

$$\int_{-\infty}^{\infty} \Phi(a + bx) \phi\left(\frac{x - \mu}{\rho}\right) \mathrm{d}x = \rho \Phi\left(\frac{a + b\mu}{\sqrt{1 + \rho^2 b^2}}\right).$$

Substituting the definition of $\underline{q}^g_{\text{SUB}}$ from Lemma 3, plugging the definitions of a_g, b_g and $\tilde{\sigma}_g$ into Equation (EC.24) and using the two Owen's formulae above completes the current step.

Step 3: An immediate corollary is that group B is under-represented if and only if $\tau_B < \tau_A$*, which by Step 2 is equivalent to Equation* (EC.23). $\square$

Proof of Part (ii). The academic merit of admitted students from group g equals

$$\begin{aligned}\mathbb{E}[q \mid Y=1, g, P_{\text{FULL}}] &= \mathbb{E}[\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{\text{FULL}}, \tilde{q}_{\text{SUB}} \geq \underline{q}^g_{\text{SUB}}, g, P_{\text{FULL}}] \\ &= \frac{\pi_g}{\tau_g C}\int_{\tilde{q}^*_{\text{FULL}}}^{\infty} \tilde{q}\Phi\left(\frac{\frac{\mu\sigma^{-2}+\tilde{q}\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}} - \underline{q}^g_{\text{SUB}}}{\frac{\sqrt{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}\sqrt{1+\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}}}\right)\phi\left(\frac{\tilde{q}-\mu}{\tilde{\sigma}_g}\right)\mathrm{d}\tilde{q}.\end{aligned}$$

By Equation (10,011.1) in Owen (1980), we get that

$$\begin{aligned}\int x\Phi(a+bx)\phi\left(\frac{x-\mu}{\rho}\right)\mathrm{d}x =& \frac{\rho^2 b}{\sqrt{1+\rho^2b^2}}\phi\left(\frac{a+b\mu}{\sqrt{1+\rho^2b^2}}\right)\Phi\left(x\sqrt{1+\rho^2b^2}+\frac{b\rho(a+b\mu)}{\sqrt{1+\rho^2b^2}}\right) \\ &-\rho^2\Phi(a+b\mu+b\rho x)\phi(x)+\mu\rho\int\Phi(a+bx)\phi\left(\frac{x-\mu}{\rho}\right)\mathrm{d}x.\end{aligned}$$

Observe that the last integral simplifies because

$$\frac{\mu\pi_g\tilde{\sigma}_g}{\tau_g C}\int_{\tilde{q}^*_{\text{FULL}}}^{\infty}\Phi(a_g+b_g\tilde{q})\phi\left(\frac{\tilde{q}-\mu}{\tilde{\sigma}_g}\right)\mathrm{d}\tilde{q} = \frac{\mu\tau_g}{\tau_g} = \mu.$$

For the first and second term, we find that

$$-\frac{\tilde{\sigma}_g^2 b_g}{\sqrt{1+\tilde{\sigma}_g^2b_g^2}}\phi\left(\frac{a_g+b_g\mu}{\sqrt{1+\tilde{\sigma}_g^2b_g^2}}\right)\Phi\left(\tilde{q}^*_{\text{FULL}}\sqrt{1+\tilde{\sigma}_g^2b_g^2}+\frac{b_g\tilde{\sigma}_g(a_g+b_g\mu)}{\sqrt{1+\tilde{\sigma}_g^2b_g^2}}\right)+\tilde{\sigma}_g^2\Phi(a_g+b_g\mu+b_g\tilde{\sigma}_g\tilde{q}^*_{\text{FULL}})\phi(\tilde{q}^*_{\text{FULL}}).$$

Putting everything together, we get that

$$\begin{aligned}\mathbb{E}[q \mid Y=1, g, P_{\text{FULL}}] =& \mu - \frac{\tilde{\sigma}_g^2 b_g}{\tau_g\sqrt{1+\tilde{\sigma}_g^2b_g^2}}\phi\left(\frac{a_g+b_g\mu}{\sqrt{1+\tilde{\sigma}_g^2b_g^2}}\right)\Phi\left(\tilde{q}^*_{\text{FULL}}\sqrt{1+\tilde{\sigma}_g^2b_g^2}+\frac{b_g\tilde{\sigma}_g(a_g+b_g\mu)}{\sqrt{1+\tilde{\sigma}_g^2b_g^2}}\right) \\ &+\Phi(a_g+b_g\mu+b_g\tilde{\sigma}_g\tilde{q}^*_{\text{FULL}})\frac{\tilde{\sigma}_g^2\phi(\tilde{q}^*_{\text{FULL}})}{\tau_g} \\ =&\lambda(a_g, b_g, \tilde{\sigma}_g, \tau_g).\end{aligned}$$

Requiring that $\mathbb{E}[q \mid Y=1, A, P_{\text{FULL}}] > \mathbb{E}[q \mid Y=1, B, P_{\text{FULL}}]$ concludes the proof. $\square$

Because testing is costly, admissions outcomes reflect both the informativeness of the test and other $K-1$ features and the cost-to-valuation ratio c_g/v. Low diversity can occur either because group B students self-select out of the test at higher rates (due to higher costs), or are admitted at lower rates even if they apply similarly (due to low feature informativeness). On the other hand, unlike exogenous barriers, student incentives can improve outcomes: higher-skilled students in both groups are more likely to take the test and apply (see Figure EC.2).

Overall, Figure EC.5 shows how test informativeness and test costs interact to determine academic merit, diversity, and individual fairness. When test costs are high for group B, both academic merit and diversity decline—an effect that is amplified when the test is *more informative* (lower conditional variance), up to a point. In such cases, more group B students self-select out of testing, exacerbating these outcomes. Intuitively, when feature informativeness is increased, group B students near the previous decision boundary have a lower admissions probability, because they can no longer can "get lucky" with a higher test score.

EC.4.6. Two schools (Proofs from Section 5.2.2 and 5.3.1)

Student decisions. In a two-school setting with policies $\mathbf{P} = (P^1, P^2)$, students' decisions to take the test and thus apply to test-requiring schools are determined per case as follows:

$$\alpha(\boldsymbol{\theta}_{\text{SUB}}, g; \mathbf{P}) = \arg\max_{\alpha \in \{0,1\}} h(\alpha, \boldsymbol{\theta}_{\text{SUB}}, g; \mathbf{P}), \tag{EC.25}$$

where

$$h(\alpha, \boldsymbol{\theta}_{\text{SUB}}, g; \mathbf{P}) = \begin{cases} \alpha\big(v_1 \Pr\big(Y_1 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^1_{\text{FULL}}\big) - c_g\big) + v_2 \Pr\big(Y_1 = 0 \cap Y_2 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^2_{\text{SUB}}\big), & \mathbf{P} = (P^1_{\text{FULL}}, P^2_{\text{SUB}}), \\ \alpha\big(v_1 \Pr\big(Y_1 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^1_{\text{FULL}}\big) + v_2 \Pr\big(Y_1 = 0 \cap Y_2 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^2_{\text{FULL}}\big) - c_g\big), & \mathbf{P} = (P^1_{\text{FULL}}, P^2_{\text{FULL}}), \\ v_1 \Pr\big(Y_1 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^1_{\text{SUB}}\big) + \alpha\big(v_2 \Pr\big(Y_1 = 0 \cap Y_2 = 1 \mid \boldsymbol{\theta}_{\text{SUB}}, g, P^2_{\text{FULL}}\big) - c_g\big), & \mathbf{P} = (P^1_{\text{SUB}}, P^2_{\text{FULL}}), \\ 0, & \mathbf{P} = (P^1_{\text{SUB}}, P^2_{\text{SUB}}). \end{cases}$$

Schools' selection policies. Recall that $Y_i(\tilde{q}_{S_i}; \mathbf{P})$ denotes the selection policy of school J_i. For brevity, we also define the indicator function

$$\chi_i(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) = 1 + (\alpha^*(\boldsymbol{\theta}_{\text{SUB}}, g; \mathbf{P}) - 1) \cdot \mathbf{1}\{P^i = P_{\text{FULL}}\}),$$

which takes value 1 in two cases: either when school J_i does not require the test (P^i_{SUB}) or school J_i requires the test and a student in group g with features $\boldsymbol{\theta}_{\text{SUB}}$ takes the test (P^i_{FULL}).

At equilibrium, given the student preference for J_1 over J_2, the more preferred school, J_1, picks students first. In particular, school J_1 optimizes the academic merit of its admitted class as follows:

$$\begin{aligned}\max_{Y_1} &\sum_g \pi_g \, \mathbb{E}_{\boldsymbol{\theta}_{S_1}}[\tilde{q}(\theta_{S_1}, g) \cdot \chi_1(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) \cdot Y_1(\tilde{q}(\boldsymbol{\theta}_{S_1}, g); \mathbf{P}) \mid g, P^1] \\ \text{s.t. } &\sum_g \pi_g \, \mathbb{E}_{\boldsymbol{\theta}_{S_1}}[\chi_1(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) \cdot Y_1(\tilde{q}(\boldsymbol{\theta}_{S_1}, g); \mathbf{P}) \mid g, P^1] \le C_1.\end{aligned} \tag{EC.26}$$

Similarly, J_2 optimizes academic merit by selecting among the students who either did not apply to J_1 at all (if J_1 requires the test) or applied but did not get admitted, i.e.,

$$\begin{aligned}\max_{Y_2} &\sum_g \pi_g \, \mathbb{E}_{\boldsymbol{\theta}_{S_2}}[\tilde{q}(\theta_{S_2}, g) \cdot \chi_2(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) \cdot Y_2(\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g); \mathbf{P}) \mid \chi_1(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) \cdot Y_1(\tilde{q}(\boldsymbol{\theta}_{S_i}, g); \mathbf{P}) = 0, g, P^2] \\ \text{s.t. } &\sum_g \pi_g \, \mathbb{E}_{\boldsymbol{\theta}_{S_2}}[\chi_2(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) \cdot Y_2(\tilde{q}(\boldsymbol{\theta}_{\text{SUB}}, g); \mathbf{P}) \mid \chi_1(\boldsymbol{\theta}_{\text{SUB}}, g, \mathbf{P}) \cdot Y_1(\tilde{q}(\boldsymbol{\theta}_{S_i}, g); \mathbf{P}) = 0, g, P^2] \le C_2.\end{aligned} \tag{EC.27}$$

Two-school equilibria. Given testing policies $\mathbf{P}$ and capacities C_i, we say that a triple $(\alpha^*, \mathbf{Y}^*, \mathbf{P})$ constitutes an *equilibrium* if: (i) for all $\boldsymbol{\theta}_{\text{SUB}} \in \mathbb{R}^{K-1}$ and $g \in \{A, B\}$, $\alpha^*(\boldsymbol{\theta}_{\text{SUB}}, g; \mathbf{P}) = \arg\max_{\alpha \in \{0,1\}} h(\alpha, \boldsymbol{\theta}_{\text{SUB}}, g; \mathbf{P})$; and (ii) for all $\boldsymbol{\theta}_{\text{FULL}} \in \mathbb{R}^K$ and $g \in \{A, B\}$, $Y_i^*(\boldsymbol{\theta}_{S_i}, g; \mathbf{P}) = \mathbf{1}\{\tilde{q}(\boldsymbol{\theta}_{S_i}, g) \ge \tilde{q}^*_{i,S_i}\}$, where $\tilde{q}^*_{i,S_i}$ is the corresponding solutions to Equation (EC.26) and Equation (EC.27).

Note that, as in the single school case, we can focus on admission strategies of the form $Y_i^*(\tilde{q}_{S_i}; \mathbf{P})$. In Lemma EC.12 below we formalize that each such Y_i preserves its threshold-based form.

Lemma EC.12. *At an equilibrium $(\alpha^*, \mathbf{Y}^*)$, each school J_i's selection policy $Y_i^*(\tilde{q}_{S_i}; \mathbf{P})$, $i \in \{1, 2\}$, takes a threshold form, i.e., there exists a threshold $\tilde{q}^*_{i,S_i}$ such that $Y_i^*(\tilde{q}_{S_i}; \mathbf{P}) = \mathbf{1}\{\tilde{q}_S \ge \tilde{q}^*_{i,S_i}\}$ where $\tilde{q}^*_{1,S_1}$, $\tilde{q}^*_{2,S_2}$, are the solutions to Equation* (EC.26) *and Equation* (EC.27), *respectively.*

Proof. We provide the proof for $\mathbf{P} = (P^1_{\text{FULL}}, P^2_{\text{SUB}})$; the remaining cases are analogous. Since $v_1 > v_2$ and school J_2 uses P^2_{SUB}, all students who apply to J_1 also apply to J_2 but not vice versa. All students have incentive to apply to J_2.

We begin with school J_1. Note that every student with $\alpha^*(\tilde{q}_{\text{SUB}}, g) = 1$ admitted to J_1 will accept the offer since $v_1 > v_2$. Therefore J_1 can pick any student as long as the student has applied to J_1. Let G_1 denote the CDF of all students with skill estimate $\tilde{q}_{\text{FULL}}$ who apply to school J_1 at equilibrium.

We show that J_1 admits the top mass $C_1^- \triangleq \min\{C_1, \int_{-\infty}^{\infty} \sum_g \pi_g a^*(\tilde{q}_{\text{SUB}}, g; \mathbf{P})\, \mathrm{d}\tilde{q}_{\text{SUB}}\}$ with the highest skill estimates $\tilde{q}_{\text{FULL}}$. I.e., $Y_1^*(\tilde{q}_{\text{FULL}}; \mathbf{P}) = \mathbf{1}\{\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{1,\text{FULL}}\}$, where $\tilde{q}^*_{1,\text{FULL}}$ satisfies

$$1 - G_1(\tilde{q}^*_{1,\text{FULL}}) = C_1^-.$$

First note that any other threshold-based policy is infeasible or suboptimal. This is because Y_1^* either admits all applicants (in the case where $C_1^- < C_1$) or the capacity constraint in Equation (EC.26) binds. Next, consider any feasible selection policy Y_1 and observe that under any Y_1 the academic merit objective in Equation (EC.26) can be written as

$$\int_{-\infty}^{\infty} \tilde{q}_{\text{FULL}} Y_1(\tilde{q}_{\text{FULL}}; \mathbf{P})\, \mathrm{d}G_1(\tilde{q}_{\text{FULL}}) = \int_0^1 G_1^{-1}(s) Y_1(G_1^{-1}(s); \mathbf{P})\, \mathrm{d}s,$$

which is trivially convex in Y_1 and supermodular. By the threshold form of Y_1^*, Y_1^* weakly majorizes any other feasible selection policy Y_1. Thus, by the Fan-Lorentz inequality (Fan and Lorentz 1954), it follows that

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{q}_{\text{FULL}} Y_1(\tilde{q}_{\text{FULL}}; \mathbf{P})\, \mathrm{d}G_1(\tilde{q}_{\text{FULL}}) &= \int_0^1 G_1^{-1}(s) Y_1(G_1^{-1}(s); \mathbf{P})\, \mathrm{d}s \\
&\leq \int_0^1 G_1^{-1}(s) Y_1^*(G_1^{-1}(s); \mathbf{P})\, \mathrm{d}s \\
&= \int_{-\infty}^{\infty} \tilde{q}_{\text{FULL}} Y_1^*(\tilde{q}_{\text{FULL}}; \mathbf{P})\, \mathrm{d}G_1(\tilde{q}_{\text{FULL}}),
\end{aligned}$$

thus Y_1^* is optimal. □

PROPOSITION EC.2 **(Proposition 2)**. *Consider the setting with two schools defined above with testing policies* $\mathbf{P} = (P^1_{\text{FULL}}, P^2_{\text{SUB}})$. *Then, there exists a unique equilibrium* $(\alpha^*, \mathbf{Y}^*)$ *with the following properties:*

(i) School J_i*'s selection policy* Y_i^* *takes a threshold form:* $Y_i^*(\tilde{q}_{S_i}; \mathbf{P}) = \mathbf{1}\{\tilde{q}_{S_i} \geq \tilde{q}^*_{i,S_i}\}$ *where* $\tilde{q}^*_{1,\text{FULL}}$, $\tilde{q}^*_{2,\text{SUB}}$, *are the solutions to* (EC.26), (EC.27), *respectively.*

(ii) Students in group g *take the test and apply to school* J_1, *if and only if one of the following conditions holds:*

1) either $\tilde{q}^*_{2,\text{SUB}} > \tilde{q}_{\text{SUB}} \geq \underline{q}^g_l$ *where*

$$\underline{q}^g_l = \tilde{q}^*_{1,\text{FULL}} - \Phi^{-1}\left(1 - \frac{c_g}{v_1}\right)\left(\frac{\sigma_{gK}^{-2}}{\sigma^{-2} + \sum_{k \in \text{FULL}} \sigma_{gk}^{-2}}\right)\sqrt{\sigma_{gK}^2 + \frac{1}{\sigma^{-2} + \sum_{k \in \text{SUB}} \sigma_{gk}^{-2}}}; \tag{EC.28}$$

2) or $\tilde{q}_{\text{SUB}} \geq \max\{\underline{q}_h^g, \tilde{q}^*_{2,\text{SUB}}\}$*, where*

$$\underline{q}_h^g = \tilde{q}^*_{1,\text{FULL}} - \Phi^{-1}\left(1 - \frac{c_g}{v_1 - v_2}\right)\left(\frac{\sigma_{gK}^{-2}}{\sigma^{-2} + \sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}\right)\sqrt{\sigma_{gK}^2 + \frac{1}{\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}. \tag{EC.29}$$

Furthermore, $\underline{q}_l^g < \underline{q}_h^g$ *for both groups* $g \in \{A, B\}$.

(iii) The fraction of students in group g *who have* $\tilde{q}_{\text{SUB}} > q_a > \underline{q}_l^g$ *and get admitted to school* J_1 *equals*

$$D_g(\hat{a}_g(q_a)) \triangleq \pi_g\tilde{\sigma}_g\Phi\left(\frac{\hat{a}_g(q_a) + b_g\mu}{\sqrt{1+\tilde{\sigma}_g^2 b_g^2}}\right) - \pi_g\tilde{\sigma}_g\Phi_2\left(\frac{\hat{a}_g(q_a) + b_g\mu}{\sqrt{1+\tilde{\sigma}_g^2 b_g^2}}, \frac{\tilde{q}^*_{1,\text{FULL}} - \mu}{\tilde{\sigma}_g}; -\frac{\tilde{\sigma}_g b_g}{\sqrt{1+\tilde{\sigma}_g^2 b_g^2}}\right),$$

where b_B *and* $\tilde{\sigma}_B$ *are defined as in Proposition EC.1, and*

$$\hat{a}_g(q_a) = \frac{\mu\sigma^{-2} - q_a}{\frac{\sqrt{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}\sqrt{1 + \frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma_{gk}^{-2}}}}.$$

Conditional on $\tilde{q}^*_{2,\text{SUB}} > \underline{q}_h^B$*, school* J_1 *is more diverse than* J_2 *if and only if*

$$\frac{D_B(\hat{\alpha}_B(\underline{q}_l^B))}{\Phi\left((\mu - \tilde{q}^*_{2,\text{SUB}})/\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\right) - D_B(\hat{\alpha}_B(\tilde{q}^*_{2,\text{SUB}}))} > \frac{C_1}{C_2}.$$

Otherwise, school J_1 *is more diverse than* J_2 *if and only if*

$$\frac{D_B(\hat{a}_B(\underline{q}_l^B)) - D_B(\hat{a}_B(\underline{q}_h^B)) + D_B(\hat{a}_B(\tilde{q}^*_{2,\text{SUB}})) + \tilde{\sigma}_B\left(\Phi\left(\frac{\hat{a}_B(\underline{q}_h^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) - \Phi\left(\frac{\hat{a}_B(\tilde{q}^*_{2,\text{SUB}})+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\right)}{\Phi\left((\mu - \tilde{q}^*_{2,\text{SUB}})/\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\right) - D_B(\hat{\alpha}_B(\underline{q}_h^B))} > \frac{C_1}{C_2}.$$

(iv) There exist instances of the model parameters such that school J_1 *achieves lower academic merit for group* g *than* J_2*. In particular, assume that* $\tilde{q}^*_{2,\text{SUB}} > \underline{q}_h^g$*. Then,* J_1 *achieves lower academic merit for group* g *than* J_2 *if and only if*

$$\lambda(a_g(\tilde{q}^*_{1,\text{FULL}}), b_g, \tilde{\sigma}_g, \tau_g) < \kappa(a_g(\tilde{q}^*_{2,\text{SUB}}), b_g, \tilde{\sigma}_g, \tau_g),$$

where

$$\kappa(a'_g(q), b'_g, \tilde{\sigma}_g, \tau_g) \triangleq \frac{\tilde{\sigma}_g^2 b'_g}{\tau_g\sqrt{1+\tilde{\sigma}_g^2(b'_g)^2}}\phi\left(\frac{a'_g(q) + b'_g\mu}{\sqrt{1+\tilde{\sigma}_g^2(b'_g)^2}}\right)\Phi\left(\tilde{q}^*_{1,\text{FULL}}\sqrt{1+\tilde{\sigma}_g^2(b'_g)^2} + \frac{b'_g\tilde{\sigma}_g(a'_g(q)+b'_g\mu)}{\sqrt{1+\tilde{\sigma}_g^2(b'_g)^2}}\right)$$
$$- \Phi\left(a'_g(q) + b'_g\mu + b'_g\tilde{\sigma}_g\tilde{q}^*_{1,\text{FULL}}\right)\frac{\tilde{\sigma}_g^2\phi(q)}{\tau_g} + \frac{\mu(1-\tau_g)}{\tau_g},$$

$$a'_g(q) = \frac{\mu\sigma^{-2} - q\left(\sigma^{-2} + \sum_{k\in\text{SUB}}\sigma_{gk}^{-2}\right)}{\sqrt{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}, b'_g = \sqrt{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}.$$

Proof of Part (i). The result was already proved in Lemma EC.12. $\square$

Proof of Part (ii). At equilibrium, all students apply to the test-free school J_2. By Part (i), only students with $\tilde{q}_{\text{SUB}} > \tilde{q}^*_{2,\text{SUB}}$ get accepted. Thus, we have two separate cases:

- *Students who get rejected by* J_2: Students in group g with $\tilde{q}_{\text{SUB}} < \tilde{q}^*_{2,\text{SUB}}$ decide to take the test (and apply to J_1) if and only if

$$v_1\,\mathbb{P}(\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{1,\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P^1_{\text{FULL}}) - c_g \geq 0,$$

  thus the problem reduces to the single-school setting. By Lemma 3, the above condition translates to $\tilde{q}^*_{2,\text{SUB}} > \tilde{q}_{\text{SUB}} \geq \underline{q}^g_l$, where Equation (EC.28) follows analogously to Equation (7) for $v = v_1$.
- *Students who get accepted to* J_2: Students in group g with $\tilde{q}_{\text{SUB}} \geq \tilde{q}^*_{2,\text{SUB}}$ decide to take the test if and only if

$$\begin{aligned} &v_1\,\mathbb{P}(\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{1,\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P^1_{\text{FULL}}) + v_2\,\mathbb{P}(\tilde{q}_{\text{FULL}} < \tilde{q}^*_{1,\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P^1_{\text{FULL}}) - c_g \geq v_2 \\ \Leftrightarrow\, &\mathbb{P}(\tilde{q}_{\text{FULL}} \geq \tilde{q}^*_{1,\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P^1_{\text{FULL}}) \geq \frac{c_g}{v_1 - v_2} \\ \Leftrightarrow\, &\tilde{q}_{\text{SUB}} \geq \underline{q}^g_h, \end{aligned} \tag{EC.30}$$

  where $\underline{q}^g_h$ in Equation (EC.29) follows similarly to Equation (7) by replacing v with $v_1 - v_2$.

The property that $\underline{q}^g_l < \underline{q}^g_h$, $g \in \{A, B\}$, follows directly from comparing Equation (EC.29) to Equation (EC.28) and using that $0 < v_1 - v_2 < v_1$.

Finally, note that the equilibrium described by Parts (i) and (ii) is unique. This follows using arguments similar to Lemma EC.11. $\square$

Proof of Part (iii). Part (ii), together with the assumption that $\tilde{q}^*_{2,\text{SUB}} > \underline{q}^B_h$, implies that students in B apply to J_1 if and only if $\tilde{q}_{\text{SUB}} \geq \underline{q}^B_l$. Thus, we can apply Step 2 in Part (i) of Proposition EC.1 and find that diversity at school J_1 equals

$$\begin{aligned} \tau^1_B &= \frac{\pi_B \tilde{\sigma}_B}{C_1}\Phi\left(\frac{a_B(\tilde{q}^*_{1,\text{FULL}}) + b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) - \frac{\pi_B\tilde{\sigma}_B}{C_1}\Phi_2\left(\frac{a_B(\tilde{q}^*_{1,\text{FULL}}) + b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{1,\text{FULL}} - \mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) \\ &= \frac{\pi_B \tilde{\sigma}_B}{C_1}\Phi\left(\frac{\hat{a}_B(\underline{q}^B_l) + b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) - \frac{\pi_B\tilde{\sigma}_B}{C_1}\Phi_2\left(\frac{\hat{a}_B(\underline{q}^B_l) + b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{1,\text{FULL}} - \mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) \\ &= D_B(\hat{a}_B(\underline{q}^B_l)) \end{aligned} \tag{EC.31}$$

Next we find the diversity level τ_B^2 at J_2. The total mass of students from group B who are eligible for acceptance at J_2 is $\pi_B \Phi\left((\mu - \tilde{q}^*_{2,\text{SUB}})/\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\right)$. However, only students who do not get admitted to J_1 actually enroll in J_2. Thus,

$$\tau_B^2 = \frac{\pi_B}{C_2}\left(\Phi\left((\mu - \tilde{q}^*_{2,\text{SUB}})/\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\right)\right)$$

Requiring that $\tau_B^1 > \tau_B^2$ gives us the condition in the statement and thus concludes the proof for the case $\tilde{q}^*_{2,\text{SUB}} > \underline{q}_h^B$.

For the complementary case where $\underline{q}_h^B \geq \tilde{q}^*_{2,\text{SUB}} > \underline{q}_l^B$, we need to take into account that students with $\tilde{q}_{\text{SUB}} \in [\tilde{q}^*_{2,\text{SUB}}, \underline{q}_h^B)$ apply and get admitted only to J_2. Formally, applying Equation (10,010.1) in Owen (1980), we find that

$$\begin{aligned}
\tau_B^1 = \frac{\pi_B}{C_1}&\left(\tilde{\sigma}_B\Phi\left(\frac{\hat{a}_B(\underline{q}_l^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) - \tilde{\sigma}_B\Phi_2\left(\frac{\hat{a}_B(\underline{q}_l^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{1,\text{FULL}}-\mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\right.\\
&+\tilde{\sigma}_B\Phi_2\left(\frac{\hat{a}_B(\underline{q}_h^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{1,\text{FULL}}-\mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\\
&\left.-\tilde{\sigma}_B\Phi_2\left(\frac{\hat{a}_B(\tilde{q}^*_{2,\text{SUB}})+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{1,\text{FULL}}-\mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\right)\\
&= \frac{1}{C_1}\left(D_B(\hat{a}_B(\underline{q}_l^B)) - D_B(\hat{a}_B(\underline{q}_h^B)) + D_B(\hat{a}_B(\tilde{q}^*_{2,\text{SUB}}))\right)\\
&+\frac{\tilde{\sigma}_B}{C_1}\left(\Phi\left(\frac{\hat{a}_B(\underline{q}_h^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right) - \Phi\left(\frac{\hat{a}_B(\tilde{q}^*_{2,\text{SUB}})+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\right).
\end{aligned}$$

For school J_2, we have that

$$\begin{aligned}
\tau_B^2 =& \frac{1}{C_2}\left(\pi_B\Phi\left((\mu - \tilde{q}^*_{2,\text{SUB}})/\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\right) - \tilde{\sigma}_B\Phi\left(\frac{a_B(\underline{q}_h^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\right.\\
&\left.+\tilde{\sigma}_B\Phi_2\left(\frac{a_B(\underline{q}_h^B)+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}, \frac{\tilde{q}^*_{1,\text{FULL}}-\mu}{\tilde{\sigma}_B}; -\frac{\tilde{\sigma}_B b_B}{\sqrt{1+\tilde{\sigma}_B^2 b_B^2}}\right)\right)\\
=& \frac{\pi_B}{C_2}\left(\Phi\left((\mu - \tilde{q}^*_{2,\text{SUB}})/\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{Bk}^{-2}}}\right) - D_B(\hat{\alpha}_B(\underline{q}_h^B))\right)
\end{aligned}$$

where the last two terms correspond to the students with $\tilde{q}_{\text{SUB}} \geq \underline{q}_h^B$ that were admitted by J_1. Requiring that $\tau_B^1 > \tau_B^2$ gives the result. $\square$

Proof of Part (iv). As in Part (iii), if $\tilde{q}^*_{2,\text{SUB}} > \underline{q}_h^g$, then students in g apply to J_1 if and only if $\tilde{q}_{\text{SUB}} > \underline{q}_l^g$. However, only a fraction of them (equal to C_1) will get admitted. In the right panel of Figure 5, this mass of admitted students is depicted in yellow. From Proposition EC.1, it follows that the academic merit of group g in the admitted class at J_1 is

$$\lambda(a_g, b_g, \tilde{\sigma}_g, \tau_g).$$

The academic merit of the admitted class in J_2 equals the expected skill of the students with $\tilde{q}_{\text{SUB}} > \tilde{q}^*_{2,\text{SUB}}$ who do not get admitted to J_1 (this is depicted in purple in Figure 5). Mathematically, we can find the merit of the admitted class to J_2 as the expected skill of all the students with $\tilde{q}_{\text{SUB}} > \tilde{q}^*_{2,\text{SUB}}$ who have $\tilde{q}_{\text{FULL}} < \tilde{q}^*_1$. Similarly to the proof of Part (ii) in Proposition EC.1, we find that

$$\mathbb{E}[q \mid Y_2 = 1, Y_1 = 0, g, P^2_{\text{SUB}}] = \frac{1}{\tau_g} \int_{-\infty}^{\tilde{q}^*_{1,\text{FULL}}} \tilde{q} \Phi \left(\frac{\frac{\mu\sigma^{-2} + \tilde{q}\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}} - \tilde{q}^*_{2,\text{SUB}}}{\frac{\sqrt{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}} \right) \phi\left(\frac{\tilde{q} - \mu}{\tilde{\sigma}_g}\right) \mathrm{d}\tilde{q},$$

which, using Equation (10,011.1) in Owen (1980), simplifies to κ as given in the statement of the proposition. $\square$

Proposition EC.3 **(Dropping tests with strategic students: Academic merit)**. *Consider two schools, J_1 and J_2, both of which initially follow test-based policies $\mathbf{P} = (P_{\text{FULL}}, P_{\text{FULL}})$. When schools optimize for academic merit only, the following statements hold:*

(i) Let $\tilde{q}^*_{1,\text{SUB}}$ be the solution to (EC.5) for $P^1 = P_{\text{SUB}}$. School J_1 drops the test if and only if (10) in Theorem 3 holds. Furthermore, conditional on J_1 dropping the test, then J_2 keeps the test if and only if (11) and (12) in Theorem 3 hold.

(ii) Let $\tilde{q}^*_{2,\text{SUB}}$ be the admission threshold of J_2 under $\mathbf{P} = (P_{\text{FULL}}, P_{\text{SUB}})$ as in Proposition 2. School J_2 drops the test, while school J_1 keeps the test, if and only if (13) and (14) in Theorem 3 hold.

(iii) There exist functions $\underline{c}_A, \underline{c}_B : \mathbb{R}_+ \to \mathbb{R}_+$ such that neither school wants to drop the test if and only if $c_A \le \underline{c}_A(c_B)$ and $c_B \le \underline{c}_B(c_A)$.

Proof of part (i). Under $\mathbf{P} = (P_{\text{FULL}}, P_{\text{FULL}})$, students must take the test to be eligible to apply to schools J_1 and J_2. A student with skill estimate $\tilde{q}_{\text{SUB}}$ takes the test if and only if

$$v_1 \, \mathbb{P}(\tilde{q}_{\text{FULL}} \ge \tilde{q}^*_{1,\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, \mathbf{P}) + v_2 \, \mathbb{P}(\tilde{q}^*_{1,\text{FULL}} > \tilde{q}_{\text{FULL}} \ge \tilde{q}^*_{2,\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, \mathbf{P}) - c_g \ge 0.$$

Recall from (EC.20) that $\tilde{q}_{\text{FULL}} \mid \tilde{q}_{\text{SUB}}, g, P_{\text{FULL}} \sim \mathcal{N}(\tilde{q}_{\text{SUB}}, \tilde{\rho}_g^2)$. Therefore, the above inequality becomes

$$v_1 - c_g + (v_2 - v_1)\Phi\left(\frac{\tilde{q}^*_{1,\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_g}\right) - v_2\Phi\left(\frac{\tilde{q}^*_{2,\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_g}\right) \geq 0. \tag{EC.32}$$

Observe that, for each group g, there exists a threshold $\underline{q}^g_{\text{FULL,FULL}}$ such that students from group g take the test if and only $\tilde{q}_{\text{SUB}} \geq \underline{q}^g_{\text{FULL,FULL}}$. This follows from the fact that the LHS of (EC.32) is strictly increasing in $\tilde{q}_{\text{SUB}}$ since Φ is strictly increasing and $v_2 < v_1$. The thresholds $\underline{q}^g_{\text{FULL,FULL}}$, $\tilde{q}^*_{1,\text{FULL}}$, $\tilde{q}^*_{2,\text{FULL}}$ are the solution to the following system of equations:

$$\underline{q}^g_{\text{FULL,FULL}} = q^*_{1,\text{FULL}} - \tilde{\rho}_g\Phi^{-1}\left(\frac{v_1 - c_g}{v_1 - v_2} - \frac{v_2}{v_1 - v_2}\Phi\left(\frac{\tilde{q}^*_{2,\text{FULL}} - \underline{q}^g_{\text{FULL,FULL}}}{\tilde{\rho}_g}\right)\right) \tag{EC.33}$$

$$C_1 = \sum_{g\in\{A,B\}} \int_{\underline{q}^g_{\text{FULL,FULL}}}^{\infty} \left(1 - \Phi\left(\frac{\tilde{q}^*_{1,\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_g}\right)\right)\phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}}\right)\mathrm{d}\tilde{q}_{\text{SUB}} \tag{EC.34}$$

$$C_1 + C_2 = \sum_{g\in\{A,B\}} \int_{\underline{q}^g_{\text{FULL,FULL}}}^{\infty} \left(\Phi\left(\frac{\tilde{q}^*_{1,\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_g}\right) - \Phi\left(\frac{\tilde{q}^*_{2,\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_g}\right)\right)\phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}}\right)\mathrm{d}\tilde{q}_{\text{SUB}}. \tag{EC.35}$$

where we used Lemma 1 in the last two equations.

Next, we study when school J_1 has the incentive to drop the test. If it does, all students apply to J_1 since there is no test cost and $v_1 > v_2$. Thus, similar to Theorem 1, school J_1 incurs an information loss (due to one missing feature) leading to academic merit decrease, but has the incentive to drop the test only if it now has access to higher-skilled candidates on average compared to its previous policy $P^1 = P_{\text{SUB}}$.

The new admission threshold $\tilde{q}^*_{1,\text{SUB}}$ that school J_1 uses is the solution to (EC.5). The proof is by case analysis.

Case (1): $\tilde{q}^*_{1,\text{SUB}} < \underline{q}^g_{\text{FULL,FULL}}$, for both $g \in \{A,B\}$. This case must be ruled out by our main assumption that C_1, C_2 are small enough such that both schools fill in their capacity. If case (i) were true, then school J_1 would not be able to fill their capacity under $\mathbf{P} = (P_{\text{FULL}}, P_{\text{FULL}})$, since the LHS in (EC.5) would be larger than the RHS in (EC.35).

Case (2): $\tilde{q}^*_{1,\text{SUB}} \geq \underline{q}^g_{\text{FULL,FULL}}$, for both $g \in \{A,B\}$. Under $P^1 = P_{\text{SUB}}$, school J_1 admits only students with $\tilde{q}_{\text{SUB}} \geq \tilde{q}^*_{1,\text{SUB}}$. Since $\tilde{q}^*_{1,\text{SUB}} \geq \underline{q}^g_{\text{FULL,FULL}}$, every student who can be admitted by

J_1 under $P^1 = P_{\text{SUB}}$ is already contained in the set of test-taking students under $P^1 = P_{\text{FULL}}$. Thus, dropping the test does not expand the pool of students from which J_1 can select its class; the relevant applicant pool is identical under $P^1 = P_{\text{FULL}}$ and $P^1 = P_{\text{SUB}}$.

Given this fixed pool and capacity C_1, the only difference between the two policies is the information used to rank applicants. Under $P^1 = P_{\text{FULL}}$, school J_1 uses the more informative signal $\tilde{q}_{\text{FULL}}$, whereas under $P^1 = P_{\text{SUB}}$ it relies only on the coarser signal $\tilde{q}_{\text{SUB}}$. By Theorem 2, when capacity is fixed and there are no access barriers, P_{FULL} yields strictly higher expected academic merit than P_{SUB}. Consequently, J_1 has no incentive to drop the test in this case, if J_2 also requires it.

Case (3): $\tilde{q}^*_{1,\text{SUB}} < \underline{q}^g_{\text{FULL,FULL}}$, for exactly one fixed group $g^* \in \{A, B\}$. First, observe that this condition is equivalent to the second part of (10). Thus, school J_1 will expand its pool of group g^* students by dropping the test. However, it might not necessarily improve its academic merit due to information loss. School J_1 improves academic merit by dropping the test if and only if

$$\sum_{g\in\{A,B\}} \pi_g \int_{\underline{q}^g_{\text{FULL,FULL}}}^{\infty} \left(\int_{\tilde{q}^*_{1,\text{FULL}}}^{\infty} \tilde{q}_{\text{FULL}} \phi\left(\frac{\tilde{q}_{\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_g}\right) \mathrm{d}\tilde{q}_{\text{FULL}} \right) \phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}} \right) \mathrm{d}\tilde{q}_{\text{SUB}}$$

$$< \sum_{g\in\{A,B\}} \pi_g \int_{\tilde{q}^*_{1,\text{SUB}}}^{\infty} \tilde{q}_{\text{SUB}} \phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}} \right) \mathrm{d}\tilde{q}_{\text{SUB}}.$$

Since $\underline{q}^{*g}_{\text{FULL,FULL}}$ is strictly increasing in c_{g^*} (see (EC.33)), the LHS is strictly decreasing in c_{g^*} given fixed capacity C_1. Note that for $c_{g^*} = 0$, the LHS is larger than the RHS; for $c_{g^*} \to v_2$, the RHS becomes larger than the LHS, since no students from group g^* take the test. Due to the continuity of the LHS in c_{g^*}, the intermediate value theorem implies that there exists a $\hat{c}_{g^*}$ such that the above inequality holds for all $c_{g^*} \geq \hat{c}_{g^*}$.

Finally, we study when school J_2 has the incentive to drop the test given that school J_1 has dropped the test. In particular, school J_1 drops the test only under case (3). Under case (3), school J_1 admits all test-taking students from group g^* under $(P_{\text{FULL}}, P_{\text{FULL}})$, i.e., students with $\tilde{q}_{\text{SUB}} \geq \underline{q}^{g^*}_{\text{FULL,FULL}}$. Furthermore, under $(P_{\text{SUB}}, P_{\text{FULL}})$, the pool of test-taking students decreases since no students from group g^* apply to J_2, while for the other group $g' \neq g^*$, $\underline{q}^{g'}_{\text{SUB,FULL}} > \underline{q}^{g'}_{\text{FULL,FULL}}$, where

$$\underline{q}^{g'}_{\text{SUB,FULL}} = \tilde{q}'_{2,\text{FULL}} - \tilde{\rho}_{g'} \Phi^{-1}(1 - c_{g'}/v_2) \tag{EC.36}$$

and $\tilde{q}'_{2,\text{FULL}}$ is school J_2's admission threshold under $(P_{\text{SUB}}, P_{\text{FULL}})$.

It is now possible that the mass of test-taking students from g' falls below C_2. Thus, to ensure that J_2 keeps the test, two conditions must hold: $M_{g'}(P_{\text{SUB}}, P_{\text{FULL}}) \geq C_2$ (i.e., condition (11)) and the expected merit of group g' under $P^2 = P_{\text{FULL}}$ is higher than the expected merit of the two admitted groups under $P^2 = P_{\text{SUB}}$, i.e.,

$$\pi_{g'} \int_{\underline{q}^{g}_{\text{SUB,FULL}}}^{\tilde{q}^*_{1,\text{SUB}}} \left(\int_{\tilde{q}^*_{2,\text{FULL}}}^{\infty} \tilde{q}_{\text{FULL}} \phi\left(\frac{\tilde{q}_{\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_{g'}}\right) \mathrm{d}\tilde{q}_{\text{FULL}} \right) \phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma \sqrt{\frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}} \right) \mathrm{d}\tilde{q}_{\text{SUB}}$$

$$> \sum_{g\in\{A,B\}} \pi_g \int_{\tilde{q}^*_{2,\text{SUB}}}^{q^*_{1,\text{SUB}}} \tilde{q}_{\text{SUB}} \phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma \sqrt{\frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}} \right) \mathrm{d}\tilde{q}_{\text{SUB}},$$

which is equivalent to (12). $\square$

Proof of Part (ii). Fixing $P^1 = P_{\text{FULL}}$, school J_2 wants to drop the test if and only if this expands its pool of high-skilled students. Similarly to case (3) in Part (i), this is equivalent to the first two conditions in (13). At the same time, school J_1 wants to keep the test if and only if either it does not expand its pool by dropping the test or the academic merit of the admitted class decreases after dropping the test. The former is similar to case (2) in part (i) and occurs if and only if the first condition in (14) holds. The latter holds if and only if the first condition in (14) does not hold but

$$\sum_{g\in\{A,B\}} \pi_g \int_{\underline{q}^{g}_{\text{FULL,FULL}}}^{\infty} \left(\int_{\tilde{q}^*_{1,\text{FULL}}}^{\infty} \tilde{q}_{\text{FULL}} \phi\left(\frac{\tilde{q}_{\text{FULL}} - \tilde{q}_{\text{SUB}}}{\tilde{\rho}_{g}}\right) \mathrm{d}\tilde{q}_{\text{FULL}} \right) \phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma \sqrt{\frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}} \right) \mathrm{d}\tilde{q}_{\text{SUB}}$$

$$< \sum_{g\in\{A,B\}} \pi_g \int_{\tilde{q}^*_{1,\text{SUB}}}^{\infty} \tilde{q}_{\text{SUB}} \phi\left(\frac{\tilde{q}_{\text{SUB}} - \mu}{\sigma \sqrt{\frac{\sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}{\sigma^{-2} + \sum_{k\in\text{SUB}} \sigma_{gk}^{-2}}}} \right) \mathrm{d}\tilde{q}_{\text{SUB}}.$$

Similarly to case (3) in Part (i), the last inequality holds if and only if $c_g \leq \hat{c}''_g$. $\square$

Proof of Part (iii). By Part (i), school J_1 wants to keep the test in Cases (1) and (2). It also wants to keep the test in Case (3) when (10) does not hold. Equivalently, by the continuity and monotonicity of M_g in $c_g, c_{g'}$, if J_1 wants to keep the test conditional on $P^2 = P_{\text{FULL}}$, there must exist continuous functions $\underline{c}^1_g : \mathbb{R} \to$ such that $c_A \leq \underline{c}^1_A(c_B)$ and

$c_B \leq \underline{c}_B^1(c_A)$. Similarly, conditional on $P^1 = P_{\text{FULL}}$, and using a similar argument to Case (3) in Part (i) and the case of school J_1 above, there must exist $\underline{c}_g^2 : \mathbb{R} \to$ such that school J_2 wants to keep the test if and only if $c_A \leq \underline{c}_A^2(c_B)$ and $c_B \leq \underline{c}_B^2(c_A)$.

Finally, $(P_{\text{FULL}}, P_{\text{FULL}})$ is an equilibrium if and only if both schools prefer to keep the test. This holds if and only if

$$c_A \leq \underline{c}_A(c_B) \quad \text{and} \quad c_B \leq \underline{c}_B(c_A),$$

where $\underline{c}_A := \min\{\underline{c}_A^1, \underline{c}_A^2\}$ and $\underline{c}_B := \min\{\underline{c}_B^1, \underline{c}_B^2\}$. $\square$

Theorem 3 **(Academic merit and two-school equilibria with strategic students)**. *Suppose that each school chooses a policy to maximize academic merit. Under school policies $\mathbf{P}$, let $M_g(\mathbf{P})$ denote the mass of test-taking students from group g and let $L_g^i(\mathbf{P})$ denote the academic merit of admitted students from group g to school J_i (see Electronic Companion Section EC.4.6 for definitions). Define also*

$$K_g(q) \triangleq 1 - \Phi\left(\frac{q-\mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}}}\right)$$

to be the mass in group g with skill estimates $\tilde{q}_{\text{SUB}} \geq q$. Then, the following hold:

(i) Policy $(P_{\text{SUB}}, P_{\text{FULL}})$ is an equilibrium if and only if

$$c_g \geq \hat{c}_g \text{ and } M_g(P_{\text{FULL}}, P_{\text{FULL}}) < K_g(\tilde{q}^*_{1,\text{SUB}}) \tag{10}$$

for exactly one group $g \in \{A, B\}$ and some threshold $\hat{c}_g > 0$, while for $g' \neq g$, it holds

$$M_{g'}(P_{\text{SUB}}, P_{\text{FULL}}) \geq C_2 \tag{11}$$

$$\textit{and} \qquad \pi_{g'} L_{g'}^2(P_{\text{SUB}}, P_{\text{FULL}}) > \pi_A L_A^2(P_{\text{SUB}}, P_{\text{SUB}}) + \pi_B L_B^2(P_{\text{SUB}}, P_{\text{SUB}}). \tag{12}$$

(ii) Policy $(P_{\text{FULL}}, P_{\text{SUB}})$ is an equilibrium if and only if

$$c_g \geq \hat{c}'_g, \quad M_g(P_{\text{FULL}}, P_{\text{FULL}}) < K_g(\tilde{q}^*_{2,\text{SUB}}) \tag{13}$$

$$\textit{and either } M_g(P_{\text{FULL}}, P_{\text{SUB}}) > K_g(\tilde{q}^*_{1,\text{SUB}}) \;\; \textit{or} \;\; c_g \leq \hat{c}''_g, \tag{14}$$

for exactly one group $g \in \{A, B\}$ and thresholds $\hat{c}''_g > \hat{c}'_g > 0$.

(iii) There exist functions $\underline{c}_A, \underline{c}_B : \mathbb{R}_+ \to \mathbb{R}_+$ such that policy $(P_{\text{FULL}}, P_{\text{FULL}})$ is an equilibrium if and only if $c_A \leq \underline{c}_A(c_B)$ and $c_B \leq \underline{c}_B(c_A)$.

Proof. The proof follows from Proposition EC.3. $\square$

PROPOSITION EC.4 **(Dropping tests with strategic students: Diversity).** *Consider two schools, J_1 and J_2, both of which initially follow test-based policies $\mathbf{P}=(P_{\text{FULL}},P_{\text{FULL}})$. When schools optimize for diversity only, school J_1 drops the test if and only if*

$$\frac{\pi\tilde{\sigma}_B}{C_1}\left(\Phi\left(\frac{a_B+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2b_B^2}}\right)-\Phi_2\left(\frac{a_B+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2b_B^2}},\frac{\tilde{q}^*_{1,\text{FULL}}-\mu}{\tilde{\sigma}_B};-\frac{\tilde{\sigma}_Bb_B}{\sqrt{1+\tilde{\sigma}_B^2b_B^2}}\right)\right)<1-\Phi\left(\Phi^{-1}(1-C_1)\sqrt{(1-\pi)\frac{\tilde{\sigma}_{A,\text{SUB}}}{\tilde{\sigma}_{B,\text{SUB}}}+\pi}\right),$$

where $\tilde{\sigma}_{g,\text{SUB}}:=\sigma\frac{\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma_{gk}^{-2}}$ and $a_B=a_B(\tilde{q}^*_{1,\text{FULL}})$, b_B, $\tilde{\sigma}_B$ are defined similarly to Proposition EC.1.

Proof. Under $P^1=P_{\text{FULL}}$, similarly to (EC.31), the diversity level at J_1 equals

$$\tau_B^1(P_{\text{FULL}},P_{\text{FULL}})=\frac{\pi\tilde{\sigma}_B}{C_1}\Phi\left(\frac{a_B+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2b_B^2}}\right)-\frac{\pi\tilde{\sigma}_B}{C_1}\Phi_2\left(\frac{a_B+b_B\mu}{\sqrt{1+\tilde{\sigma}_B^2b_B^2}},\frac{\tilde{q}^*_{1,\text{FULL}}-\mu}{\tilde{\sigma}_B};-\frac{\tilde{\sigma}_Bb_B}{\sqrt{1+\tilde{\sigma}_B^2b_B^2}}\right).$$

If J_1 drops the test, then all students apply to J_1. Since $v_1>v_2$, J_1 does not compete with J_2, thus under $P^1=P_{\text{SUB}}$, the diversity level at J_1 equals

$$\tau_B^1(P_{\text{SUB}},P_{\text{FULL}})=1-\Phi\left(\Phi^{-1}(1-C)\sqrt{(1-\pi)\frac{\tilde{\sigma}_{A,\text{SUB}}}{\tilde{\sigma}_{B,\text{SUB}}}+\pi}\right)$$

as in the single school setting without barriers (see the proof of Part (i) in Theorem 2). Requiring that $\tau_B^1(P_{\text{SUB}},P_{\text{FULL}})>\tau_B^1(P_{\text{FULL}},P_{\text{FULL}})$ gives the statement. $\square$

EC.5. 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 EC.9.

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

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

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

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

and for K features,

$$\int_\infty^\infty \Pi_k f(\theta_k|g,q)\,\mathrm{d}F(q) = \frac{e^{\left(-\frac{\sum_{k=1}^K\left[\left(\mu+\mu_{gk}-\theta_k\right)^2\sigma^{-2}\sigma_{gk}^{-2}\right]+\sum_{k\neq\ell}\left[\left((\mu_{\ell g}-\theta_\ell)-(\mu_{gk}-\theta_k)\right)^2\sigma_{\ell g}^{-2}\sigma_{gk}^{-2}\right]}{2\left(\sigma^{-2}+\sum_{k=1}^K\sigma_{gk}^{-2}\right)}\right)}}{2\,(1-\pi)^{K/2}\sigma\left(\Pi_k\sigma_{gk}\right)\sqrt{\sigma^{-2}+\sum_{k=1}^K\sigma_{gk}^{-2}}} \tag{EC.37}$$

Thus, we have

$$\begin{aligned} w(\boldsymbol{\theta},g) &\triangleq \frac{f(\boldsymbol{\theta}\,|g)\,\mathbb{P}(g)}{f(\boldsymbol{\theta})} = \frac{\int_\infty^\infty \Pi_k f(\theta_k|g,q)\,\mathrm{d}F(q)\,\mathbb{P}(g)}{f(\boldsymbol{\theta})} \\ &\propto \frac{\mathbb{P}(g)\exp\left\{\left(-\frac{\sum_{k=1}^K\left[\left(\mu+\mu_{gk}-\theta_k\right)^2\sigma^{-2}\sigma_{gk}^{-2}\right]+\sum_{k\neq\ell}\left[\left((\mu_{\ell g}-\theta_\ell)-(\mu_{gk}-\theta_k)\right)^2\sigma_{\ell g}^{-2}\sigma_{gk}^{-2}\right]}{2\left(\sigma^{-2}+\sum_{k=1}^K\sigma_{gk}^{-2}\right)}\right)\right\}}{\left[\Pi_k\sigma_{gk}\right]\sqrt{\sigma^{-2}+\sum_{k=1}^K\sigma_{gk}^{-2}}} \end{aligned}$$

Derivation for equation (EC.37). We explicitly show the algebra for $K=1$ and $K=2$ features, and the pattern continues for K features.

For one feature θ_1:

$$
\begin{aligned}
w(\theta_1, g) &\triangleq \frac{f(\theta_1|g)\,\mathbb{P}(g)}{f(\theta_1)} = \frac{\int_{\infty}^{\infty} f(\theta_1|g,q)\,\mathrm{d}F(q)\,\mathbb{P}(g)}{f(\theta_1)} \\
&= \frac{\frac{1}{\sqrt{2(1-\pi)(\sigma^2+\sigma_{g1}^2)}} \exp\left[-\frac{(\mu+\mu_{g1}-\theta_1)^2}{2(\sigma^2+\sigma_{g1}^2)}\right]\mathbb{P}(g)}{f(\theta_1)} = \frac{\frac{1}{\sqrt{\sigma^2+\sigma_{g1}^2}} \exp\left[-\frac{(\mu+\mu_{g1}-\theta_1)^2}{2(\sigma^2+\sigma_{g1}^2)}\right]\mathbb{P}(g)}{\sum_g \left[\frac{1}{\sqrt{\sigma^2+\sigma_{g1}^2}} \exp\left[-\frac{(\mu+\mu_{g1}-\theta_1)^2}{2(\sigma^2+\sigma_{g1}^2)}\right]\mathbb{P}(g)\right]}.
\end{aligned}
$$

For two features θ_1, θ_2:

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

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

EC.6. General distributions

Extended model. We extend the model from Section 2 to non-Normal distributions. In the current setting, each candidate is characterized by a (latent) *true skill* q drawn from a distribution F^0 with support $Q=[\underline{q},\overline{q}]$ and mean μ.[18] We assume that F^0 is common for both social groups.

For each candidate, the school has access to K observable *features* $\boldsymbol{\theta}=(\theta_k)_{k=1}^K$. Throughout this section, we thus assume that the school uses policy P_{FULL} and omit it from the notation.

Conditional on the true skill level q and group g, feature θ_k is independently drawn from a distribution $F_{q,g}^k$. Let $\Theta_k=[\underline{\theta}_k,\overline{\theta}_k]$ be the support of each feature θ_k. We assume that the distributions F^0, $\{F_{q,g}^k\}_{k=1}^K$ are common knowledge. Without loss of generality and for the sake of simplicity, we further assume that F^0, $\{F_{q,g}^k\}_{k=1}^K$ are continuous (although being measurable would suffice).

[18] Formally, we assume that there exists a probability space $(Q,\mathcal{F},\mathbb{P})$ on which q is defined.

At an aggregate level per group g, the information structure $\left(\times_{k=1}^{K}\Theta_k, F^0, \{F^k_{q,g}\}_{k=1}^{K}\right)$ induces a *skill estimate* distribution, $\hat{F}_g$, for candidates in group g, i.e., $\tilde{q} \mid g \sim \hat{F}_g$, where $\tilde{q}(\boldsymbol{\theta}, g) \triangleq \mathbb{E}[q \mid \boldsymbol{\theta}, g]$ as in the main model. We also let $\hat{F} = (1-\pi)\hat{F}_A + \pi\hat{F}_B$.

Preliminaries. We will need the following technical terms and properties.

Definition EC.1 (**Blackwell (1953)**). $\{\Pi_q, q \in Q\}$ is *sufficient* for $\{\Pi'_q, q \in Q\}$ if there exists a transformation $T(x, dy)$ such that for all $q \in Q$, $\Pi'_q(\cdot) = \int_X T(x, \cdot)\Pi_q(dx)$.

Lemma EC.13 **(Eckwert and Zilcha (2004), Zhang (2009))**. *The following statements are equivalent:*

- $\{\Pi_q, q \in Q\}$ *is sufficient for* $\{\Pi'_q, q \in Q\}$;
- *The distribution of posteriors* $\langle \Pi'_q \rangle$ *second-order stochastically dominates* $\langle \Pi_q \rangle$.

Lemma EC.14 **(Gentzkow and Kamenica (2016))**. *If the distribution of posteriors* $\langle \Pi_q \rangle$ *is a mean-preserving spread of* $\langle \Pi'_q \rangle$*, then the posterior mean distribution under* $\langle \Pi_q \rangle$ *is a mean-preserving spread of the posterior mean distribution under* $\langle \Pi'_q \rangle$.

Lemma EC.15. *Let* X, Y *be two random variable with equal means* $\mathbb{E}[X] = \mathbb{E}[Y]$*, support* $[\underline{q}, \overline{q}]$*, and CDFs* F *and* G*, respectively. Then, the following are equivalent:*

(i) $Y \succ_{SSD} X$;

(ii) X *is a mean-preserving spread of* Y;

(iii) $\int_{\underline{q}}^{\overline{q}} u(y)\,\mathrm{d}G(y) \geq \int_{\underline{q}}^{\overline{q}} u(x)\,\mathrm{d}F(x)$ *for every weakly increasing concave function* $u : [\underline{q}, \overline{q}] \to \mathbb{R}$.

Lemma EC.16. *Let* X *and* Y *be two random variables with support* $[\underline{q}, \overline{q}]$ *and CDFs* F *and* G*, such that* Y *is a mean preserving of* X*. Then,* F *crosses* G *exactly once at a point* $q_+ \in [\underline{q}, \overline{q}]$*. For* $q \in [\underline{q}, q_+)$*,* $F(q) > G(q)$ *whereas for* $q \in (q_+, \underline{q}]$*,* $F(q) < G(q)$.

Generalizing Proposition 1. We are finally ready to prove a generalized version of Proposition 1.

Proposition EC.5. *Suppose that* $(\tilde{q} \mid A) \prec_{SSD} (\tilde{q} \mid B)$ *with crossing point* q_+*. Consider a school that uses admissions policy* P_{FULL}*. Then,* $(\tilde{q} \mid A) \prec_{SSD} (\tilde{q} \mid B)$ *is equivalent to each of the following conditions.*

(i) Diversity*: Group* B *is under-represented if and only if* $C < 1 - \hat{F}(q_+)$;

(ii) Academic merit*: For any capacity* C*, the policy achieves worse academic merit for admitted students from group* B.

Furthermore, suppose that $\{\times_{k=1}^{K} F_{q,A}^{k}, q \in Q\}$ is sufficient for $\{\times_{k=1}^{K} F_{q,B}^{k}, q \in Q\}$. Equivalently,

(iii) Individual fairness*: there exists a threshold $\hat{q}$ such that $I(q; P_S) > 0$ if and only if $q > \hat{q}$.*

Proof. We prove each part separately.

Proof of part (i). Let $\tilde{q}^*$ denote the optimal acceptance threshold as given by the following equation:

$$(1-\pi)\hat{F}_A(\tilde{q}^*) + \pi \hat{F}_B(\tilde{q}^*) = 1 - C \Longleftrightarrow \hat{F}(\tilde{q}^*) = 1 - C.$$

Therefore, for $C < 1 - \hat{F}(q_+)$, it holds that $\tilde{q}^* > q_+$, and vice versa. Thus, part (i) follows directly from Lemma EC.16.

Proof of part (ii). Part (ii) follows from the equivalence between (i) and (iii) in Lemma EC.15 where we consider $u(x)$ to be the linear function $u(x) = x$.

Proof of part (iii). By Lemma EC.13, sufficiency equivalently guarantees that the posterior distribution $\{\times_{k=1}^{K} F_{q,A}^{k}, q \in Q\}$ second-order stochastically dominates $\{\times_{k=1}^{K} F_{q,B}^{k}, q \in Q\}$. By Lemma EC.16, this immediately translates to the following property:

$$\mathbb{P}[Y = 1 \mid q, A] = \mathbb{P}[\tilde{q} \geq \tilde{q}^* \mid q, A] > \mathbb{P}[Y = 1 \mid q, B] = \mathbb{P}[q \geq \tilde{q}^* \mid q, B]$$

if and only $q > \tilde{q}^*$, where $\tilde{q}^*$ is the optimal acceptance threshold corresponding to some capacity C. $\square$

Note that an analog of the above proposition can also be obtained for any subset of features S.

EC.7. Affirmative action

Schools often have an additional lever in their admissions policies: whether or not to use affirmative action. The term *affirmative action* refers to admissions policies that partially base decisions on applicants' membership in social groups with legally protected characteristics (e.g., race, ethnicity, or gender), to promote equal opportunity as well as the educational benefits of diversity (Alon 2015).

We define affirmative action as a constraint on the fraction of students from each group. As a result, the admissions policy may use different admission thresholds for different groups. This approach is common in the literature (Fang and Moro 2011) and a proxy of the practices adopted by universities. However, due to the recent lawsuit against Harvard (Hartocollis 2019) and the Supreme Court decision in 2023 (Saul 2023), the legal framework around such affirmative action is restrictive. Explicit, predetermined *racial* quotas are generally illegal, as is (newly) broad consideration of race separate from individuals' contexts; conversely, University of Texas admits students using a high school-based quota system (The University of Texas 2019).

From a theoretical standpoint, the class of affirmative-action policies is interesting because it generates a Pareto frontier between the academic merit and diversity objectives. A fully Bayesian school, using group information when forming skill estimates but then accepting students with the highest skill estimates regardless of group, would maximize academic merit. To instead maximize some weighted combination of academic merit and diversity, an optimal school (with no legal constraints) would be fully Bayesian *within* each group, ranking students within each group according to their expected true skill and then accepting the top students *from each group* to achieve some desired balance between academic merit and diversity objectives. Different weights would correspond to different fractions of students from each group, tracing out a Pareto curve.

Next, we study outcomes when schools can decide both whether to require standardized testing *and* whether to use affirmative action.

Affirmative action under a fixed testing policy. As a stylized model of affirmative action, we extend the main setup of Section 2 by introducing 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 must 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 $\tilde{q}^*_{B,S}$.[19] Note that $\tau(P_S^\tau) = \tau$, thus under affirmative action, diversity improves by definition, and group fairness holds when the target diversity level is set to $\tau = \pi$.[20] Affirmative action can be utilized on top

[19] In Proposition EC.6, 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 $\tilde{q}^*_{g,S} \geq \mu$).

[20] Proposition EC.6 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.

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 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 EC.6 **(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., $I(q; P_S^\tau) > 0$, 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). With affirmative action, the common threshold $\tilde{q}^*_S$ in Equation (EC.4) is replaced by two group-dependent thresholds, $\tilde{q}^*_{A,S}$ and $\tilde{q}^*_{B,S}$:

$$(1-\pi)\gamma_A(1 - F_{\tilde{q}|A,P_S}(\tilde{q}^*_{A,S})) = (1-\tau)C, \quad \pi\gamma_B(1 - F_{\tilde{q}|B,P_S}(\tilde{q}^*_{B,S})) = \tau C. \tag{EC.38}$$

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.

First, observe that Equation (EC.38) gives us

$$\tilde{q}^*_{A,S} = F^{-1}_{\tilde{q}|A,P_S}\left(1 - \frac{1-\tau}{(1-\pi)\gamma_A}C\right),\ \tilde{q}^*_{B,S} = F^{-1}_{\tilde{q}|B,P_S}\left(1 - \frac{\tau}{\pi\gamma_B}C\right). \tag{EC.39}$$

Since $\tau > \tau(P_S)$ and $\gamma_B \le \gamma_A \le 1$, it follows that $\tilde{q}^*_{B,S} < \tilde{q}^*_S < \tilde{q}^*_{A,S}$. Due to our assumptions that $\gamma_A \ge \frac{2(1-\tau)C}{1-\pi}$ and $\gamma_B \ge \frac{2\tau C}{\pi}$, 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 $\tau(P_S) < \tau \le \pi$, $\mathbb{P}[\tilde{q} \ge \tilde{q}^*_{A,S} \mid q, A, P^\tau_S] < \mathbb{P}[\tilde{q} \ge \tilde{q}^*_S \mid q, A, P_S]$, and $\mathbb{P}[\tilde{q} \ge \tilde{q}^*_{B,S} \mid q, B, P^\tau_S] > \mathbb{P}[\tilde{q} \ge \tilde{q}^*_S \mid q, B, P_S]$, since the distribution of $\tilde{q} \mid q, P$ remains the same under both $P \in \{P_S, P^\tau_S\}$. Consequently, $I(q; P^\tau_S) < I(q; P_S)$.

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

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

which is equivalent to

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

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

$$\sqrt{\sum_{k\in S}\sigma^{-2}_{Bk}}\sqrt{\sum_{k\in S}\sigma^{-2}_{Ak}} > \sigma^{-2}, \tag{EC.40}$$

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

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

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, $I(\hat{q}; P^\tau_S) < 0$.

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

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

To see why, observe that given the condition in Equation (EC.40), the function

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

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 this particular instance of model parameters. Consequently, we have constructed an instance of model parameters such that $I(\hat{q}; P_S) > 0 > I(\hat{q}; P_S^\tau)$ for some $\hat{q}$. Thus, such an instance exists. $\square$

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 Equation (EC.10), 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. $\square$

Dropping the test under affirmative action. We now study how test-free and test-based policies with affirmative actions compare in a setting with unequal barriers γ_g to test access. Recall that Theorem 1 (without affirmative action) shows 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. Recall that the function HR denote the hazard rate of the Normal distribution Φ, $\mathrm{HR}(z) = \frac{\phi(z)}{1-\Phi(z)}$.

PROPOSITION EC.7 **(Dropping tests under affirmative action with barriers).** *Fix group $g \in \{A, B\}$, variances σ^2_{gk}, 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., $\mathbb{E}[q \mid Y=1, g, P^\tau_{\text{FULL}}] < \mathbb{E}[q \mid Y=1, g, P^\tau_{\text{SUB}}]$, if and only if $\gamma_g \leq \hat{\gamma}_g$, where*

$$\hat{\gamma}_g = \frac{\tau_g C}{1-\Phi\left(\mathrm{HR}^{-1}\left(\frac{\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma^{-2}_{gk}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma^{-2}_{gk}}}}{\sqrt{\frac{\sum_{k\in\text{FULL}}\sigma^{-2}_{gk}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma^{-2}_{gk}}}}\mathrm{HR}(\Phi^{-1}(1-\frac{\tau_g C}{\pi_g}))\right)\right)}. \tag{EC.41}$$

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

$$t_g = \frac{\tilde{w}^*_{g,\text{FULL}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{FULL}}\sigma^{-2}_{gk}}{\sigma^{-2}+\sum_{k\in\text{FULL}}\sigma^{-2}_{gk}}}}, \quad t'_g = \frac{\tilde{q}^*_{g,\text{SUB}} - \mu}{\sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma^{-2}_{gk}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma^{-2}_{gk}}}}.$$

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

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

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

$$\mathbb{E}[q \mid Y=1, g, P^\tau_{\text{SUB}}] = \mu + \sigma\sqrt{\frac{\sum_{k\in\text{SUB}}\sigma^{-2}_{gk}}{\sigma^{-2}+\sum_{k\in\text{SUB}}\sigma^{-2}_{gk}}} \cdot \frac{\phi(t'_g)}{1-\Phi(t'_g)}.$$

To compute the threshold $\hat{\gamma}_g$, we require that $\mathbb{E}[q \mid Y=1, g, P^{\tau}_{\text{SUB}}] = \mathbb{E}[q \mid Y=1, g, P^{\tau}_{\text{FULL}}]$. 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 Equation (EC.39) 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 Equation (EC.41). Note that the expected skill level of admitted students in the test-based policy is given – due to Lemma EC.3 – by

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

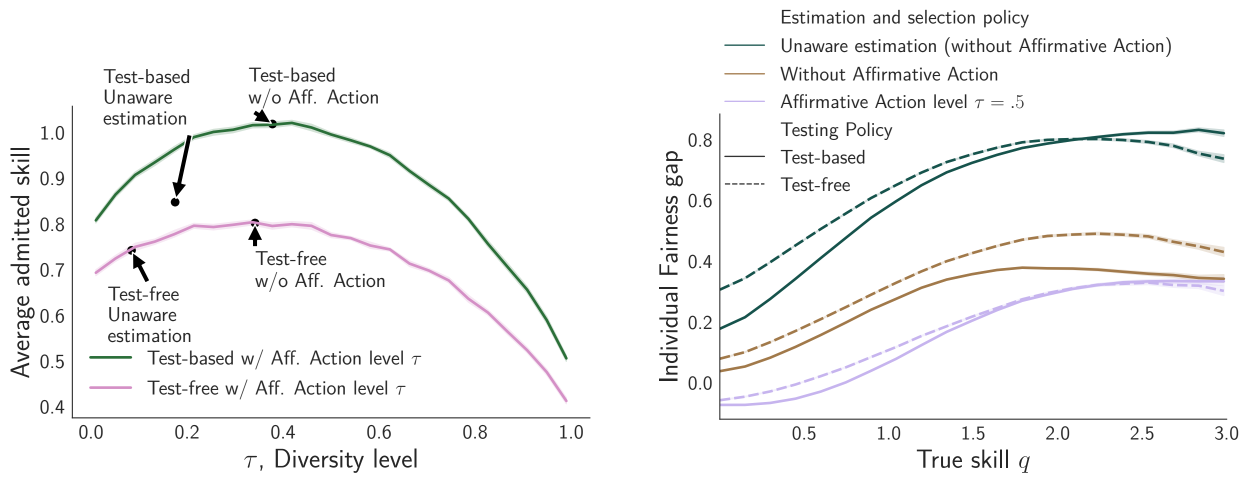
By Lemma EC.4, it follows that HR is increasing. However, $\Phi^{-1}\left(1-\frac{\tau_g C}{\pi_g\hat{\gamma}_g}\right)$ is decreasing in $\hat{\gamma}_g$. Therefore, the academic merit of g must be decreasing in $\hat{\gamma}_g$. Thus, dropping the test increases academic merit for g if and only if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

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 EC.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 Equation (EC.41) is decreasing in σ_{gK}. Consequently, $\hat{\gamma}_g$ increases as σ_{gK} increases. $\square$

Observe that 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.



(a) Average admitted skill vs. diversity level

(b) Individual fairness gap

Figure EC.9 Performance of various policies, in simulation in a setting where features are more informative for group A, and with testing barriers for group B. Group-unaware policies, analyzed in Section EC.5, are those in which the school does not use group information in its Bayesian skill estimation. Affirmative action, analyzed in Section EC.7, is defined as fixing the diversity level τ as a constraint. Figure (a) then shows results for the full range of diversity levels. Affirmative action in general improves both diversity and individual fairness, while dropping the test score has an ambiguous impact. Group-unaware policies generally perform the worst on all metrics. More generally, dropping the test has equivocal effects, depending on the objective (diversity level, academic merit, individual fairness gap) and other policies. The full parameter set can be found in Electronic Companion EC.2.1.2.

Comparing the policies in simulation. Figure EC.9 compares, for one parameter setting, our policies: with and without testing, and with and without *affirmative action* (where a fixed fraction τ of the admitted class is group B; see Section EC.7). In Figure EC.9a, the Pareto curves trace the trade-off between diversity and academic merit, for each testing policy. In this scenario, constraining each group's admitted class to be proportional to its group size (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 equivocal 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 EC.9 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 Electronic Companion EC.5.

E-Companion References

Agarwal N, Somaini P (2018) Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica* 86(2):391–444.

Agarwal N, Somaini P (2020) Revealed preference analysis of school choice models. *Annual Review of Economics* 12(1):471–501.

Ajayi K, Sidibe M (2020) School choice under imperfect information. *ERID Working Paper* (294).

Alon S (2015) *Race, Class, and Affirmative Action* (Russell Sage Foundation).

Baricz Á (2008) Mills' ratio: Monotonicity patterns and functional inequalities. *Journal of Mathematical Analysis and Applications* 340(2):1362–1370.

Blackwell D (1953) Equivalent comparisons of experiments. *The Annals of Mathematical Statistics* 265–272.

Calsamiglia C, Fu C, Güell M (2020) Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy* 128(2):642–680.

College Board (2022) California SAT Suite of Assessments Annual Report. URL `reports.collegeboard.org/media/pdf/2022-california-sat-suite-of-assessments-annual-report.pdf`.

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

Eckwert B, Zilcha I (2004) Economic implications of better information in a dynamic framework. *Economic Theory* 24(3):561–581.

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

Fan K, Lorentz GG (1954) An integral inequality. *The American Mathematical Monthly* 61(9):626–631.

Fang H, Moro A (2011) Theories of Statistical Discrimination and Affirmative Action: A Survey. *Handbook of Social Economics*, volume 1, 133–200 (Elsevier).

Gentzkow M, Kamenica E (2016) A Rothschild-Stiglitz approach to bayesian persuasion. *American Economic Review* 106(5):597–601.

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

Idoux C (2023) Integrating New York City schools: The role of admission criteria and family preferences .

Kapor AJ, Neilson CA, Zimmerman SD (2020) Heterogeneous Beliefs and School Choice Mechanisms. *American Economic Review* 110(5):1274–1315, ISSN 0002-8282, URL http://dx.doi.org/10.1257/aer.20170129.

Owen DB (1980) A table of normal integrals: A table. *Communications in Statistics-Simulation and Computation* 9(4):389–419.

Rothstein JM (2004) College performance predictions and the sat. *Journal of Econometrics* 121(1-2):297–317.

Saul S (2023) After Affirmative Action Ends. *The New York Times* URL https://www.nytimes.com/2023/06/30/us/politics/affirmative-action-college-admissions-supreme-court.html.

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

The University of Texas (2019) Top 10 percent law. URL news.utexas.edu/key-issues/top-10-percent-law/.

Tienda M, Sullivan TA (2011) *Texas Higher Education Opportunity Project* (Inter-university Consortium for Political and Social Research).

Tomkins S, Grossman J, Page L, Goel S (2023) Showing high-achieving college applicants past admissions outcomes increases undermatching. *Proceedings of the National Academy of Sciences* 120(45):e2306017120.

Zhang Z (2009) *Comparison of information structures with infinite states of nature* (The Johns Hopkins University).