Testing Away from One's Own School: Exam Location and Performance in High-Stakes Exams*

Justin Jihao Hong † Victor Lei ‡ Xuan Li §

This version: September 2025

Abstract: High-stakes exams are often administered at designated test centers, requiring many students to test in unfamiliar environments. We investigate whether such arrangements impact students' test performance and, by extension, access to educational opportunities. Using unique administrative data from China's national college entrance examination (2016 – 2018) and its random assignment of test centers, we find that students assigned to a non-home school score 0.14 standard deviations lower than classmates testing at their home school, and they are 3.8 percentage points less likely to be admitted to college. Mechanism analysis points to unfamiliar environments as a primary driver in our context, while longer travel distances are also associated with poorer performance. We further highlight substantial inequality implications: the penalty is especially severe for low-achieving students and those from disadvantaged backgrounds. As test centers are predominantly located in high-performing schools, such ostensibly neutral policies may unintentionally exacerbate existing achievement gaps between privileged and less privileged groups. A back-of-the-envelope calculation suggests that exam location accounts for over 7.6% of the observed performance gap between students from test-center and non-test-center schools.

Keywords: High-Stakes Test, Exam Location, Cognitive Performance, Disparity

JEL: D90, H75, I23, I24, I28

*We thank George Bulman, Yiming Cao, Raymond Fisman, Joshua Goodman, Le Kang, Kevin Lang, Daniele Paserman, and participants at ACEM Doctoral Students Forum (SJTU 2025) and NSD Applied Micro PhD Conference (PKU 2025) for their helpful comments. All errors are our own.

†Hong Kong University of Science and Technology (GZ) (email: hjihao@bu.edu)

‡University of Hong Kong (email: victorlei.econ@gmail.com)

§Boston University (corresponding author, email: xuanli@bu.edu)

1 Introduction

Most students take standardized tests before graduating from high school. These high-stakes exams play a crucial role in college admissions and can have lasting effects on long-term outcomes. A key organizational characteristic of these tests is that they are typically administered at designated test centers, which are usually higher-quality schools. As a result, students — particularly those from disadvantaged backgrounds — often need to take the exam in unfamiliar, non-home schools. While prior research has shown that such testing arrangements can influence the likelihood of taking the exam (Bulman, 2015), much less is known about their impact on test performance. Psychological evidence suggests that unfamiliar environments can trigger situational anxiety and impair cognitive functioning, particularly during critical, high-stakes moments in life (Smith, 1979; Nejati, 2023). This raises important concerns that such testing policies may unintentionally exacerbate existing educational inequalities.

A central challenge in identifying the causal effect of exam location lies in self-selection and the endogenous choice of where to take the exam. For example, students taking the Scholastic Assessment Test (SAT) can choose whether to register and select their preferred test center at the time of registration. In this paper, we address the challenge by leveraging the random assignment of test centers in China's national college entrance examination (NCEE), drawing on administrative data on the full population of high schoolers in a representative county. Nearly all high school graduates in China register for the NCEE as university admissions are based solely on NCEE scores. To reduce administrative costs and deter cheating, local education authorities pool all NCEE candidates within their jurisdiction (typically at the county level) after registration and *randomly* assign them to designated test centers — usually local high schools. Accordingly, students are unaware of their test center assignment at the time of registration, and the random assignment introduces exogenous variation in exam locations. Crucially for our identification, this assignment mechanism generates *within-class* variation for students from high schools designated as test centers: while some remain at their own school to take the exam, others are required to take in different schools (see Figure A1 for a visual illustration). This setting enables comparisons between otherwise similar students who take the exam in different testing environments.

Using data on the entire population of approximately 11,000 students who took the NCEE in our sampled county from 2016 to 2018, we first validate the random assignment system. We show a student's assignment to a non-home school is not correlated with demographic characteristics, including gender, age, and socioeconomic status (SES).

After confirming the random assignment, we proceed to estimate the impact of testing at a non-home school on exam performance. We find that students assigned to a non-home school score 0.14 standard deviations lower than those taking the exam at their home school — equivalent to a 10-point reduction on the 750-point scale used in China's NCEE. The performance penalty associated with non-home location remains persistent throughout the sample period, and it is consistently observed across all high schools and among both male and female students. To put the magnitude of non-home performance decline into context, we examine its implications for students' access to future educational opportunities. Students assigned to non-home schools are ranked behind an additional 2.9% of their peers within the same year-track in their province, are 3.8 percentage points less likely to be admitted to any college, and are 0.6 percentage points less likely to gain admission to an elite college in the year of the exam.

We then discuss two major potential channels: unfamiliar environment and longer commute. We present two pieces of evidence suggesting unfamiliar environment as a significant mechanism. First, the performance penalty is primarily driven by STEM subjects, consistent with the idea that environmental factors — such as unfamiliar testing environments — can heighten stress and consume cognitive resources particularly critical for STEM performance (Beilock and Carr, 2005). Second, we find that the presence of a familiar individual (i.e., a classmate) seated nearby can mitigate the adverse effects of testing in a non-home school, providing further support for the unfamiliar environment channel. Regarding longer commute, we find that longer travel time is negatively associated with exam performance, even after controlling for individual characteristics including socioeconomic status. This suggests that increased fatigue and logistical uncertainty may exacerbate cognitive decline. However, controlling for travel time does not alter the magnitude of the effect of testing in a non-home school: in our context, test centers are relatively concentrated, and students assigned to non-home schools travel only slightly farther on average than those assigned to their own schools. While longer travel distance cannot meaningfully explain the observed performance penalty in our setting, it remains an important factor in contexts where traveling to the test center requires a significant amount of time.

Finally, we discuss the broader implications of our findings for understanding educational disparities. We highlight that the performance penalty associated with testing location is both (1) more severe for less privileged students, and (2) more likely to affect them. One, our within-class comparison reveals that the performance penalty is more pronounced among lower-performing and socioeconomically disadvantaged students, consequently impacting their admission outcomes. Although students have the option to retake the exam the following year, this response is largely limited to those from higher socioeconomic backgrounds, likely due to the substantial costs involved. Two, we show that — consistent with patterns observed in the United States and other settings — test centers are predominantly located in more developed areas, and thus closer to more advantaged groups. This implies that students from less privileged backgrounds are more likely to be assigned to non-home test centers and disproportionately bear the performance penalty. Our back-of-the-envelope calculation suggests that 7.6% of the observed performance gap between students from test-center high schools and those from non-test-center schools could be attributed to the exam location.

Our paper primarily contributes to the discussion on socioeconomic gaps in education, by highlighting an integral component of college access processes that has received little attention. The existing studies focus on students' decisions regarding whether to take or retake college entrance exams (i.e., the extensive margin) (Bulman, 2015; Goodman, 2016; Frisancho et al., 2016; Goodman, Gurantz and Smith, 2020; Kang et al., 2024). Another line of research examines how environmental or institutional factors influence students' performance gaps (i.e., the intensive margin) (Ebenstein, Lavy and Roth, 2016; Graff Zivin et al., 2020a; Park, 2022; Bond et al., 2022; Chang and Padilla-Romo, 2023; Wang, Wang and Ye, 2023; Li et al., 2024). Our context allows us to net out test environments (via test room fixed effects) and students' self-selection, enabling us to uncover the impact of exam location on test performance. We find that students who take exams at a school other than their own score 0.14 standard deviations lower than those who test at their home school, resulting in reduced educational opportunities. This effect is particularly pronounced among low-achieving students and those from disadvantaged backgrounds. Combined with the fact that high-performing high schools are more likely to serve as test centers, our results suggest that ostensibly neutral assignment policies may inadvertently widen existing achievement gaps between privileged and less privileged groups.

The study most closely related to ours is Bulman (2015), which examines inequalities in access to test centers — specifically, how the presence of a test center at a student's home high school affects their likelihood of taking a college admissions exam.[1] We further complement the analysis by focusing on the intensive margin: how taking a high-stakes exam at a non-home school, rather than at one's own school, affects test performance. We provide causal evidence in China's college entrance exam, one of the most high-stakes exams in the country. Our results suggest that exam location imposes an intensive margin penalty on top of the extensive-margin barriers documented in Bulman (2015). Given the uneven global distribution of test centers, this double penalty tied to test location carries broader implications for educational inequality and access to opportunity.

2 Background and Data

2.1 Setting: National College Entrance Examination (NCEE)

In China, the National College Entrance Examination (NCEE), or *Gaokao*, is widely regarded as one of the most demanding and consequential standardized tests in the world. It functions as the primary — and often the sole — criterion for university admission. In contrast to admissions systems like that of the United States, Chinese universities base admissions almost exclusively on *Gaokao* scores. The exam is administered only once a year, in early June, and spans two consecutive days. For most students, this single exam largely determines their access to higher education and profoundly influences their long-term career prospects (Jia and Li, 2021).

The Organization of NCEE (*Gaokao*). The Chinese government places strong emphasis on the organization of the NCEE to ensure a smooth and efficient process. The *Gaokao* registration period takes place from November 1 to November 10 of the preceding year. For instance, students who took the 2016 *Gaokao* registered between November 1 and 10, 2015. In each county, several high schools are designated as test centers. All registered students within the same county take the exam at these centers, which typically contain 20 to 40 testing rooms, each accommodating 30 students.

Most provinces randomly assign students to test rooms based solely on their academic track (i.e., science or liberal arts), ensuring that all students in a given room follow the same

[1] In the U.S., fewer than half of public high schools serve as SAT test centers, with even lower rates among schools serving low-income communities. Since the SAT's introduction in 1901, proximity to a test center has played a crucial role in determining access to college entrance exams. Bulman (2015) finds that opening a new test center increases test-taking by an average of 8.5 percent (4 percentage points) among students at the host school, with about 40 percent of these students subsequently enrolling in a four-year college.

track, without considering other factors such as registration timing. This randomization minimizes opportunities for collusion among students or between students and proctors — who are typically recruited from other counties to further reduce the risk — and reflects broader efforts to promote fairness in the examination process nationwide. Since 2016, as part of broader efforts to promote standardization — and a key component of our identification strategy — the central government has required local education authorities to randomly assign test rooms within designated administrative areas.[2] While the assignment of students to test rooms is fully randomized, whether graduating students are allowed to take the exam in their own high schools (if those schools serve as test centers) can be determined by local governments. Our context has advantages in this regard. In some provinces, all students take the exam at their own schools if those schools are designated as test centers; in others, the majority do — for example, about 90% of students remain at their own schools. By contrast, in our sample province (and several others), assignment is entirely random: students are allocated with equal probability to any test center within the designated area.[3] Therefore, among students graduating from high schools that serve as test centers in our sample province, some are randomly assigned to take the exam at their home school, while others are randomly assigned to different test centers.

Specifically, the provincial government assigns students to specific test centers, rooms, and seats one week before the exam.[4] Students are notified of their assigned test center and room several days after the randomization, and these assignments remain fixed throughout the entire examination period. Switching test centers, rooms, or seats is strictly prohibited. Every seat in the exam room is occupied, and proctors strictly enforce seating arrangements. Each assigned seat is labeled with the student's name, photo, and identification details, which are verified by proctors before the start of each exam.

Students are allowed to visit the test center the afternoon before the exam to familiarize themselves with their assigned room and nearby facilities, such as restrooms. The test rooms will be re-inspected after the students' familiarization visit. Participation in this pre-exam visit is optional. During the examination, each room is supervised by two proctors — one male

[2]See *Regulations on National College Entrance Examination Administration (2016)*.

[3]There are anecdotal reports of top high schools in our sample province lobbying the government to modify the fully random assignment process — specifically, to allow all of their students to take *Gaokao* at their own schools. While such arrangements may provide logistical benefits for the schools, they could also enhance student performance by allowing students to test in a familiar and less stressful environment.

[4]A computer program is employed to randomly assign students to test rooms. Specifically, within each county and academic track, students are allocated according to computer-generated random numbers.

and one female — who are high school teachers recruited from other counties and randomly assigned. After the exam, the papers are scanned by computer and then randomly assigned to experienced graders for evaluation in a double-blind process. Appointed by the provincial education authority, these graders convene for one week at a government-designated location to evaluate exam scripts.[5]

Notably, since the NCEE is a high-stakes exam that students spend 12 years preparing for, we find that in our county, all registered students ultimately sit for the exam. In other words, whether students take the exam is unrelated to their assigned test centers in our context.

Exam Subjects. During our sample period (2016 – 2018), our sample province adopted the National Unified Exam Paper, which is designed and administered by the National Education Examinations Authority under the Ministry of Education. This exam was used in most provinces across China, meaning that a large number of students nationwide faced the same exam structure, content, and scoring standards. The exam consists of four sections totaling 750 points: Chinese (150), Mathematics (150), English (150), and a comprehensive test (300). There are generally two tracks: the science track and liberal arts track. The comprehensive test covers physics, chemistry, and biology for science track students, and politics, history, and geography for liberal arts track students. The selection of academic track occurs at the end of Grade 10 (two years before the NCEE). Students could not choose individual subjects within their track, as subject combinations are fixed.

College Admission. Students typically receive their test scores about two weeks afterward. Upon receiving their scores, they begin formulating their college application lists. While multiple provinces may use the same version of the exam (e.g., the National Unified Exam Paper Type I), college admissions are conducted separately by province. Each province has its own admission quotas, cut-off scores, and ranking systems. As a result, students compete only against others within their own province.

Admission outcomes depend on both the number of available spots and a student's rank

[5] To enhance exam integrity, additional protocols have been implemented during the NCEE to prevent cheating: (1) while students are permitted to use the restroom during the exam, few do so due to time constraints, and those who do are accompanied and monitored by a same-sex proctor; (2) students are prohibited from wearing school uniforms to prevent identification with specific high schools; (3) students are prohibited from bringing any electronic devices into the test room; (4) all examinees are individually screened with metal detectors upon entry, and radio signal jammers are used to ensure full and effective coverage of the test rooms; (5) monitoring cameras are installed in every test room, recording the entire exam process in real time under the supervision of the provincial authorities; and (6) the nearby road will be closed to motor vehicles in order to reduce the noise level.

among provincial peers applying to the same institution. Each year, over 10 million students take the *Gaokao*, including more than 500,000 from our sample province alone. While approximately 40% gain admission to college, only a small fraction are accepted into top-tier universities, resulting in an intensely competitive environment with extraordinarily high stakes. Given the within-province competition, the *Gaokao* is highly competitive — where even a single point can significantly affect a student's rank. To illustrate this competitiveness, consider our sample province in 2018, which had approximately 200,000 science track examinees. In this context, a student with a score of 500 achieved a provincial rank of 63,000. A reduction of just one point would lead to a drop of 680 positions, disadvantaging the student relative to hundreds of peers in the college selection process.

The province we study reflects the typical structure of China's higher education system. It offers a full range of institutions, from top-tier universities (including Project 985 schools) to more general colleges, covering nearly all types of degree-granting programs. Most students in our sample choose to attend college within the province.

2.2 Data

Our data is sourced from the Bureau of Education of a county in Central China. This administrative dataset covers all students who registered for the national college entrance examination in our sample county from 2016 to 2018. It contains detailed information on each student's gender, ethnicity, age, academic track, high school and class, test center, and test room. The dataset also includes subject-level test scores and, importantly, college admission outcomes.[6]

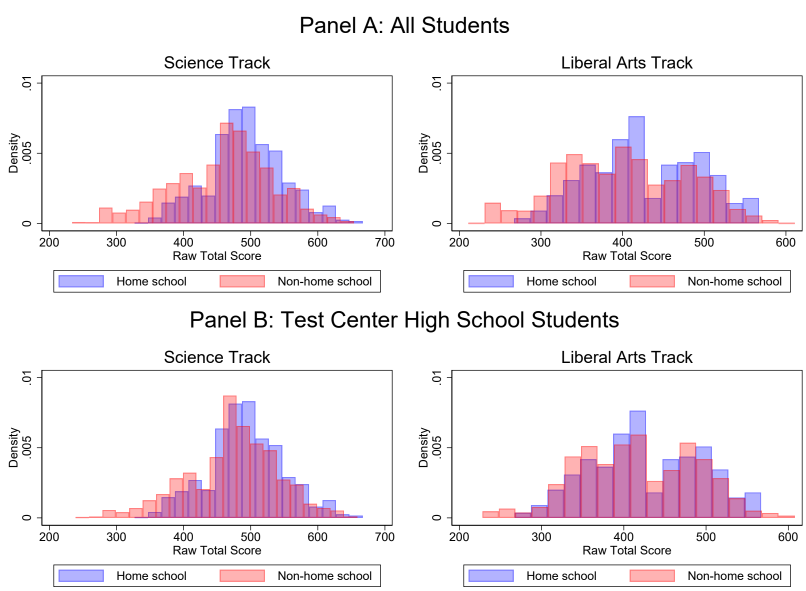
The county has a geographical size comparable to that of Houston or Greater London. Table A1 provides background information regarding the sampled county characteristics. For comparison, we also provide corresponding statistics for an average county. While our sam-

[6] We are not the first to utilize the student-level National College Entrance Exam data (Cai et al., 2019). Previous literature utilizes data from a single province or several provinces, with some even employing data from the entire country (between 1999 and 2003) (Graff Zivin et al., 2020a; Li and Qiu, 2023; Li et al., 2024; Guo, Shi and Zhang, 2024; Kang et al., 2024). However, our identification necessitates that (1) the sampled area implements random test center assignment during the sampling period, and (2) the data includes detailed administrative information, including the test center and high school. Provinces in China began implementing random test room and test center assignments relatively recently, a practice that extends beyond the sample period covered in existing literature. To the best of our knowledge, the datasets at the provincial or national scale utilized in the existing literature do not include information that can determine whether a student take the exam in his/her own high school. Our resources only allow us to access data in our sample county that meets both of these conditions simultaneously. Nevertheless, since we use data in a more recent period, it helps us to rule out some competing channels, such as cheating, as the administrative is more standardized and strict in more recent years.

pled county is less prosperous than the average, it has a comparable share of high schoolers in its population and a similar student-teacher ratio relative to other counties in the same province.

Each year, approximately 4,000 students from eleven high schools in our sample county register for the national college entrance examination. Five of these schools serve as designated test centers, each accommodating a roughly equal number of students, though some host slightly more than others.[7] Specifically, the largest test center hosts 21% of students, while the smallest hosts 18.4%. Figure A2 presents the spatial distribution of high schools and designated test centers to show their relative positions. All five test centers are centrally located within the county situated closer to the local government offices. Students from the most remote high school must travel 9 kilometers — often navigating mountainous roads by bus or electric bicycle — to reach their designated test center for the exam.

Figure 1: Distribution of Raw Total Scores



Notes: This figure shows the distribution of raw total scores by test center status (home school vs. non–home school), separately for students in the science and liberal arts tracks. Panel A includes all students, while Panel B focuses on students attending high schools that serve as test centers. The histograms are based on raw data. Blue bars represent students who took the exam at their home schools, while red bars represent those who took the exam at non-home schools. The x-axis displays raw total scores, and the y-axis indicates the density.

[7]These five high schools are the top-performing schools in the area. We provide a detailed comparison of test-center and non–test-center schools in Section 4.2.

Figure 1 presents the distribution of raw total scores for students in two academic tracks, comparing those who took the exam at their home school (in blue) versus a non-home school (in red). The left figure in Panel A shows the distribution for the Science Track. Both groups exhibit approximately bell-shaped distributions, but home-school test takers tend to cluster at slightly higher score ranges. On average, students who took the exam at their home school scored 492 points, compared to 459 points for those who tested at a non-home school, indicating a substantial performance gap. The right figure in Panel A shows displays the distribution for the Liberal Arts Track, where the score distributions are more similar across test center types. Nonetheless, home-school test takers still show a slight concentration in the upper score range, with average scores of 427 for home-school test takers and 401 for their non-home school counterparts. Overall, the figure highlights consistent performance differences by test center type: students who take the exam at their home schools tend to perform better. However, this initial comparison reflects the influence of two factors: (1) the potential performance penalty associated with the exam location, and (2) underlying differences between students from schools that serve as test centers and those from schools that do not (all of whom are non–home-school test takers).

Because our primary research question focuses on isolating the effect of exam location on performance, our analysis centers on students from the five high schools in the county that serve as test centers. These schools offer meaningful within-school or within-class variation in exam location, as each includes students who took the exam both at their home school and at a different test center. In contrast, students from schools that are not designated as test centers must take the exam elsewhere and thus exhibit no variation in home-school status. These five top high schools enroll 81% of all students in the county. Among them, 82.9% took the exam at a non-home school. The maximum distance between any two of these schools is 6.5 kilometers. Panel B of Figure 1 presents the distribution of raw total scores for students from the five high schools that serve as test centers, disaggregated by the location where they took the exam. While the score gap between home-school and non-home-school exam takers has narrowed, students who took the exam at their home school still clearly performed better.

3 Empirical Strategy

3.1 Identification

In this section, we move beyond descriptive evidence and present more rigorous analysis. Our baseline specification is defined as follows:

$$TotalScore_{ic(s)r(t)} = \beta NonHomeSchool_{ic(s)r(t)} + \pi_{c(s)} + \pi_{r(t)} + \epsilon_{ic(s)r(t)} \quad (1)$$

where student i attends high school s, belongs to high school class c, and takes the NCEE in test room r at test center t. $TotalScore_{ic(s)r(t)}$ represents student i's standardized total score on the NCEE. To ensure comparability of test scores across different years, we standardize test scores within each year and academic track (i.e., liberal arts or science) to have a mean of 0 and a standard deviation of 1. $NonHomeSchool_{ic(s)r(t)}$ is a dummy variable that is one if the student s is assigned to a test center outside their own home school (i.e., $s \neq t$).

Crucially for our identification, we introduce π_c, the high school class fixed effects, which allow us to compare students within the same class and thus absorb any *high school* $\times$ *year* $\times$ *track* differences. In China, students are typically placed into classes according to their academic track and ability at the outset of 11th grade. Each class consists of 25 to 45 students with similar academic performance, and these students attend all of their courses together with the same group of classmates. The inclusion of high school class fixed effects enhances the comparability of our sample by controlling for these class-level groupings. This approach also addresses potential selection concerns, as some schools host a larger number of exam takers, and students from these schools are more likely to take the exam at their own institution. Given the inclusion of high school class fixed effects, our analysis focuses on students from the five high schools in the county that serve as test centers (approximately 8,500 students).[8] These schools provide meaningful within-school variation in exam location, as each includes students who took the exam both at their home school and at a different test center.[9]

[8] It is important to note that all students in our analysis are first-time exam takers, as these five high schools do not admit repeat test-takers, who are typically enrolled in specialized repeat-year schools. In Section 4.1, we provide additional analysis of repeat exam takers who graduated from these five schools in the previous year and find similar results.

[9] In contrast, all students from non-designated schools are required to take the exam at an alternative location and, consequently, exhibit no variation in home-school exam status. Although these students are excluded from the main analysis for identification purposes, they constitute an important group for understanding the implications for educational inequality, which we discuss in detail in Section 4.2.

We also include test room fixed effects, π_r, which account for differences in the testing environment among students assigned to the same room, capturing factors such as temperature and the condition of school facilities. β is the coefficient of interest, capturing the performance gap attributable to assignment to non-home test centers. Standard errors are clustered at the high school class level.

3.2 Balance checks

Given the random assignment process, it is reasonable to expect that a student's assignment to a non-home school is uncorrelated with individual-level characteristics. To validate the randomness of test center assignment, we examine whether a student's placement at a non-home school is associated with individual-level characteristics.

Table A2 presents summary statistics and balance checks for demographic and background variables available in our data, comparing students taking exams at home schools and non-home schools. 49.3% of students are male, the average age is 17.6 years, 38.6% of students originate from urban areas, 81.7% are enrolled in the science track, and 52.3% live in neighborhoods with house prices above the median (i.e., ¥2500 per square meter).[10] The results show no statistically significant differences between the two groups in terms of gender, age, class leadership roles, participation in the science track, or socioeconomic status (proxied by house price). Moreover, approximately 22% of students have a classmate in the same test room, and this proportion is identical for those testing in their home school and in other schools, as all students are randomly assigned to test centers and test rooms. These findings confirm successful randomization and comparable student populations in our context.

4 Results

4.1 Main results

Table 1 presents our main results. We begin by regressing test scores on non-home school assignments, incorporating only the high school class fixed effects. This approach enables comparisons among students within the same high school class — a comparable group with similar academic backgrounds and an equal likelihood of being assigned to different test centers. We find that students who take the exam at their home school perform substantially

[10] Notably, the urban–rural classification is based on household registration and thus reflects students' places of origin. The house price indicator is based on current home addresses and serves as a better proxy for socioeconomic status.

better. In Column (2), we include test center fixed effects to account for factors specific to each center, such as location and the condition of school facilities. All high schools that served as test centers in our sample county underwent renovations between 2006 and 2008 and thus have relatively modern facilities. We again find that students who take the exam at their home school perform substantially better, with a similar magnitude of effect.

Table 1: Main Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total Score				Rank Percentile	College Admission	Elite College Admission	Exam Retake
Non-Home School	-0.136***	-0.133***	-0.141***	-0.140***	2.930***	-0.0375***	-0.00583*	0.133***
	(0.0174)	(0.0172)	(0.0186)	(0.0184)	(0.425)	(0.0103)	(0.00330)	(0.0115)
Observations	8,535	8,535	8,535	8,535	8,535	8,535	8,535	5,619
R-squared	0.486	0.487	0.512	0.520	0.550	0.465	0.215	0.100
Individual Controls	.	.	.	X	.	.	.	.
Test Center FEs	.	X	.	.	.	.	.	.
Test Room FEs	.	.	X	X	X	X	X	X
Highschool Class FEs	X	X	X	X	X	X	X	X

Notes: This table presents our main results. *Non-Home School* is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. *Total Score* represents the student's total test score, standardized by year and academic track. *Rank Percentile* measures provincial rank percentile. *College Admission* is a binary variable indicating whether student i is admitted to any college in that year. *Elite College Admission* is a binary variable indicating whether student i is admitted to an elite college in the current year. *Exam Retake* is a binary variable indicating whether student i retook the exam in the following year. Individual controls include gender, age, class leadership status, urban residency, the presence of classmates, and socioeconomic status, proxied by housing prices. Standard errors are clustered at the high school class level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Column (3) presents our preferred specification, which incorporates both high school class fixed effects and test room fixed effects. The inclusion of test room fixed effects accounts for test room-specific factors, such as floor location, student composition, proctors, and other environmental influences. The estimated coefficient remains virtually unchanged. In Column (4), we add individual controls, including gender, age, class leadership status, urban residency, the presence of classmate in the same test room, and socioeconomic status. The results remain consistent. On average, students assigned to a non-home school score 0.14 standard deviations lower than their classmates who take the exam at their home school.

In addition, Column (1) of Panel A in Table A3 reports an alternative measure of exam performance: the raw total score. Students who took the exam at a non-home school scored, on average, 10 points lower (out of 750), representing a 1.3% decrease in the total score. To put this into perspective, since each multiple-choice question typically carries 3 to 6 points, this gap is roughly equivalent to answering two to three more questions incorrectly across all six subjects.[11]

The importance of score changes is contingent on their influence on college attendance outcomes. Students typically receive their test scores about two weeks after taking the NCEE. Upon receiving their scores, they begin formulating their college application lists. College admissions are conducted separately by province. Notably, each province has its own admission quotas, cut-off scores, and ranking systems. As a result, students compete only against others within their own province. Chinese universities are grouped into different admission tiers — First, Second, and Third Tiers — based on their overall quality or prestige. In each tier, students are allowed to apply to up to five schools. Universities admit students by selecting the highest-scoring applicants who listed them as a preference, continuing until their admission quotas are filled. Once a student is admitted by a university, they are removed from the applicant pool. As a result, each student can be admitted to only one university. If their score falls short for all the universities they applied to, they will not receive an offer of admission.[12]

Since college admission outcomes in China depend more on a student's relative ranking among applicants from the same province and academic track (rather than on their absolute test scores), it is important to understand how changes in test scores translate to shifts in provincial ranking percentiles. In Column (5), we match our data to the score distribution in the province, accounting for year and academic track, and calculate the provincial rank percentiles, which measures the proportion of students with higher test scores within the same year-track. On average, the decline in test scores places affected students behind an additional 2.9% of their peers within the same year-track in the province.

Subsequently, in Column (6) of Table 1, we find that the decline in test scores translates into a meaningful reduction in immediate access to educational opportunities. In our sam-

[11] To understand how this magnitude compares to the effects of short-run environmental and psychological shocks on student performance in other contexts, consider the following examples: in Mexico, violent crimes occurring in the week prior to exams reduce female students' test scores by 0.11 standard deviations (Chang and Padilla-Romo, 2023); in Israel, a 10-unit increase in PM2.5 exposure lowers scores by 0.08 standard deviations (Ebenstein, Lavy and Roth, 2016). Moreover, Park (2022) finds that taking the Regents Exams in New York City when outdoor temperatures reach 90°F reduces performance by approximately 0.13 standard deviations compared to taking the exam at 75°F. In the Chinese context, a 5°C (9°F) increase in temperature during the national college entrance exam period reduces total test scores by 0.15 standard deviations (Graff Zivin et al., 2020b). Students taking the NCEE in China may be more sensitive to external environmental shocks such as temperature due to its exceptionally challenging nature, coupled with the intense competition and high pressure surrounding it.

[12] For any given student, typically only one of the three tiers is relevant, depending on their exam score. For example, a high-scoring student is likely to be admitted to a university in the First Tier, making any of their applications to schools in the Third Tier effectively irrelevant.

ple county, 67.6% of students are admitted to college in the year of their exam. However, those assigned to non-home schools are 3.8 percentage points less likely to be admitted — a 5.6% decrease relative to the baseline.[13] Among more than 2,000 universities in China, 39 are designated as top-tier institutions under the "985 Initiative". Column (7) shows that students assigned to non-home schools are 0.6 percentage points less likely to gain admission to one of these elite universities in the year of their exam. It is worth noting that this result should be interpreted with caution though, as admission outcomes are influenced by both test scores and students' application strategies (Li and Qiu, 2023).

Lastly, NCEE takers who believe their initial scores do not accurately reflect their true abilities or meet their expectations may choose to retake the exam the following year, even if some have already been admitted to a college in the current year (Kang et al., 2024). These retakers typically remain in their home county to prepare for and sit the exam again. Because we have data on the full population of exam takers in our sample county from 2016 to 2018, we are able to identify retakers from the 2016 and 2017 cohorts; this results in a smaller sample size for analysis. To do so, we match observations across two consecutive years using full name, exact date of birth, gender, and academic track (science or liberal arts). Individuals successfully matched to records in the subsequent year are classified as having retaken the NCEE. In our sample county, approximately 19.7% of students retook the NCEE in the following year. In the last column of Table 1, we find that students assigned to a non-home school are significantly more likely to be dissatisfied with their admission outcomes and are 13.3 percentage points more likely to retake the exam the following year.

We conduct an array of additional analyses to better characterize the non-home school penalty and its potential drivers.

Robustness. We use alternative clustering levels in Panel B of Table A3. In Column (1), standard errors are clustered at the school-year-track level, while in Column (2), they are clustered at the test room level.

In Column (3) of Panel B in Table A3, we present additional analysis of repeat exam takers who graduated from five high schools serving as test centers. A student can retake the exam the following year if they are not satisfied with their first attempt, and some may be randomly assigned to their high school alma mater for their second attempt. Because of limited within-

[13] This analysis focuses on initial admission outcomes and does not account for students who may enter college in later years through exam retakes or adult education programs.

class variation in this specification — resulting from the small number of exam retakers — we instead include only school-track-year fixed effects.[14] Additionally, we control for their test scores from their previous *Gaokao*, which serve as a baseline measure of academic ability in a high-stakes examination setting. We find similar results: these students perform better when taking the exam at their alma mater.

Table A4 presents heterogeneity results disaggregated by year, gender, and high schools. In Panel A, We demonstrate that our results are not driven by a specific year, as the non-home school performance penalty persists throughout the sample period. Furthermore, this performance penalty is gender-neutral, with both male and female students experiencing a similar effect. Panel B of Table A4 further reports results disaggregated by high school. We find that the observed performance penalty is not driven by any single school; rather, it is consistently present among students from all high schools that serve as test centers. Therefore, our results are unlikely to be explained by the notion that non-home schools have poorer facilities that negatively impact academic performance. While the magnitude of the effect varies across schools, this likely reflects differences in student composition.

Table 2: Results by Subjects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Math	Chinese	English	Comp. test (Science)	Comp. test (Liberal Arts)	Total Score	
Non-Home School	-0.205***	-0.0527**	-0.0486*	-0.153***	-0.0977*	-0.151***	-0.114*
	(0.0250)	(0.0238)	(0.0255)	(0.0205)	(0.0544)	(0.0198)	(0.0605)
Observations	8,535	8,535	8,535	6,965	1,499	6,965	1,499
R-squared	0.333	0.325	0.325	0.504	0.422	0.578	0.411
Test Room FEs	X	X	X	X	X	X	X
Highschool Class FEs	X	X	X	X	X	X	X
Academic Track	All	All	All	Science	Liberal Arts	Science	Liberal Arts

Notes: This table presents our results by subjects. $Non\text{-}Home\ School$ is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. $Total\ Score$ represents the student's total test score, standardized by year and academic track. $Chinese$ represents the student's standardized test score on Chinese. $Math$ represents the student's standardized test score on Math. $English$ represents the student's standardized test score on English. $Comp.\ test\ (Science)$ refers to the student's standardized score on the comprehensive science exam, which includes physics, chemistry, and biology. $Comp.\ test\ (Liberal\ Arts)$ refers to the standardized score on the comprehensive liberal arts exam, covering history, geography, and political science. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

[14] Repeat test takers cannot enroll in regular high schools and instead attend specialized repeat-year schools. In this specification, school-year fixed effects refer to their original high schools rather than the specialized repeat-year schools.

Potential Mechanisms. First, according to psychological studies, unfamiliar environments can impair cognitive functioning (e.g, Smith, 1979). High-stakes exams such as the *Gaokao* are inherently anxiety-inducing for most students, given their pivotal role in shaping future educational and career trajectories. When students are assigned to take the exam at a different school — an unfamiliar setting — they may experience elevated situational anxiety, a temporary state of heightened stress triggered by new environmental conditions. Such anxiety can adversely affect working memory, attention, and problem-solving abilities, all of which are essential for optimal performance — particularly in cognitively demanding subjects like mathematics. As a result, students tested outside their home school may face a psychological disadvantage. Although our high-stakes, real-world setting precludes us from providing direct neuroscientific evidence as psychologists have done in laboratory experiments, Table 2 offers suggestive evidence by examining the impact disaggregated by subject.

STEM subjects — particularly math and science — require sustained attention, multi-step reasoning, and problem-solving under pressure. As a result, any discomfort or distraction in the testing environment is more likely to impair cognitive performance in these subjects compared to those that rely more on recall or reading comprehension. We find that STEM subjects are more sensitive to exam location. Students assigned to a non-home school score 0.2 standard deviations lower in math and 0.15 standard deviations lower in the comprehensive science test (which includes physics, chemistry, and biology). In contrast, the impact is smaller for language subjects, with scores decreasing by approximately 0.05 standard deviations in Chinese and English, and 0.1 standard deviations in the comprehensive liberal arts test (which includes history, geography, and political science).[15] Although geography is typically categorized as a non-STEM subject, it includes both physical and human geography. The former often entails calculations comparable to those in STEM fields, which may partly account for the slightly larger effect observed in the comprehensive liberal arts test. Finally, Columns (6) and (7) provide suggestive evidence that students in the science track are relatively more affected by non-home test taking than those in the liberal arts track, although the difference is not statistically significant due to limited power.[16]

A complementary approach to testing the "unfamiliar environment" hypothesis is to examine whether factors that ease unfamiliarity (e.g., similar spatial layouts or the presence of

[15] The analysis of the comprehensive test has a smaller sample size because the science version is taken only by students in the science track, while the liberal arts version is taken only by students in the liberal arts track.

[16] Since the negative effect is more pronounced in STEM subjects, we further examine in Appendix A whether taking the exam at a non-home school influences students' choice of college type or major.

familiar peers) help. To this end, we study whether having a familiar individual nearby mitigates environmental unfamiliarity and reduces its adverse effects. We first examine whether the presence of a classmate in the same test room alleviates the negative effect of testing in a non-home school in Column (1) of Table 3, but we do not find evidence supporting this. This may be because the two students are seated relatively far apart, so the feeling of unfamiliarity is not significantly reduced. As shown in Figure A3, exam regulations specify that each test room must contain exactly 30 seats arranged in five columns and six rows. Therefore, to better capture the concept of mitigating environmental unfamiliarity through the presence of familiar individuals, we construct a new variable measuring whether a student has a classmate seated nearby. For example, for Seat 14, neighboring seats include 10, 11, 12, 13, 15, 22, 23, and 24. Approximately 5% of students have a classmate sitting nearby during the *Gaokao*. Results are reported in Column (2) of Table 3. We find that these students, when surrounded by a classmate, are less affected by taking the exam at a non-home school. This lends some further credibility to the "unfamiliar environment" hypothesis.[17]

It is worth noting that cheating is unlikely to be a concern in our setting. As noted in Footnote 5, during our sample period, real-time video surveillance and rigid protocol rules make the NCEE one of the most difficult exams to cheat on (Borcan, Lindahl and Mitrut, 2017). Since 2015, individuals caught facilitating mass cheating in the NCEE can face up to seven years in prison. No such cases have been reported in our sampled county. Quantitatively, given the consistency of our results across different test centers and years, it is unlikely that undetected, isolated incidents of cheating could meaningfully explain our main results. Furthermore, even if cheating does occur — which is highly unlikely — it is difficult to explain why having a classmate seated nearby in a non-home school would have a larger positive effect on performance compared to having a classmate nearby in a home school.

[17]We verify that whether a student has a classmate sitting nearby is uncorrelated with whether they are testing at their home school (due to the random assignment of seats), indicating that students at both home and non-home schools have an equal likelihood of having a nearby classmate. Therefore, while this analysis supports the "unfamiliar environment" channel, the presence of a nearby classmate is not the underlying factor accounting for the observed performance penalty associated with exam location.

Table 3: Suggestive Evidence on Mechanisms

	(1)	(2)	(3)	(4)
	Total Score			
Non-Home School	-0.143***	-0.148***	-0.126***	-0.124***
	(0.0202)	(0.0190)	(0.0182)	(0.0180)
Non-Home School × Classmate	0.00909			
	(0.0544)			
Non-Home School × Nearby Classmate		0.148*		
		(0.0821)		
Travel Time			-0.00414***	-0.00415***
			(0.000607)	(0.000612)
Observations	8,535	8,535	8,224	8,224
R-squared	0.512	0.513	0.493	0.501
Individual Controls	.	.	.	X
Test Room FEs	X	X	X	X
Highschool Class FEs	X	X	X	X

Notes: This table presents suggestive evidence on mechanisms. $Non\text{-}Home\ School$ is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. $Classmate$ is a binary variable equal to 1 if the student has at least one classmate in the same room, and 0 otherwise. $Nearby\ Classmate$ is a dummy variable equal to 1 if a classmate is seated nearby. $Travel\ Time$ measures the travel time, in minutes, from the student's residential address to the test center by e-bike. $Total\ Score$ represents the student's total test score, standardized by year and academic track. $Classmate$ and $Nearby\ Classmate$ are included in Columns (1) and (2), respectively. Standard errors are clustered at the high school class level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Second, traveling to the test center may partially contribute to the non–home-school penalty. Longer travel distances are likely associated with reduced sleep duration. The *Gaokao* begins at 9 a.m., and exam takers are required to arrive 45–60 minutes early to complete security checks and verify their identity. Many students also experience insomnia due to exam-related stress, making any additional commute time that reduces sleep particularly consequential. Consistent with this, Heissel and Norris (2018) find that school start times and sleep patterns can significantly influence student academic performance. Moreover, longer commutes may increase fatigue and logistical uncertainty, both of which can impair cognitive functioning. This effect may be especially relevant in the U.S. context, where some parents must drive long distances to drop their children off for the SAT.

To investigate this mechanism, we calculate the travel time between each student's home and their assigned test center using e-bike transportation, the most common mode of transportation in our sampled county.[18] On average, it takes a student 24.6 minutes to travel from home to the test center, with the 25th percentile at 13 minutes and the 75th percentile at 35 minutes. We include travel time in Column (3) of Table 3 and make two observations: (1) longer travel time is associated with poorer academic performance; (2) including travel time does not meaningfully alter the magnitude of the non-home school effect. This suggests that while travel time is correlated with poorer performance, it does not substantially account for the observed non-home school penalty in our context. This result may be attributed to the fact that, in our sample county (as shown in Figure A2), test schools are relatively concentrated, and most students reside throughout the central area of the county. Consequently, on average, students assigned to non-home schools travel only slightly farther than those assigned to home schools, as confirmed by the distribution of travel times by group in Figure A4. Since students living far away may have different characteristics than those living very close to the test center, Column (4) further includes individual controls such as socioeconomic status (proxied by housing price), and we find very similar results.[19]

To better understand the relationship between travel time and test performance, and to what extent it contributes to the non-home school penalty, we run an augmented version of the specification in Column (4) of Table 3. Specifically, we breaks down the continuous travel

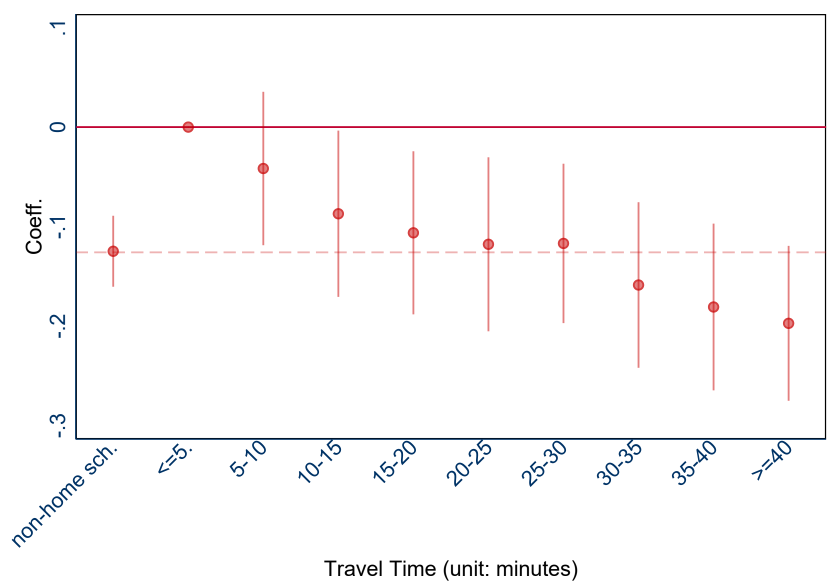
[18] The calculation is performed by a local government agent using our code, which relies on the AMap API — the Chinese equivalent of Google Maps. Due to changes in street names, however, travel time information is unavailable for around 300 students.

[19] Students residing at greater distances may mitigate the travel effect by staying in accommodations near the test center. However, in the absence of such data, our estimates are likely to be understated.

time variable by categorizing it into 5-minute intervals, using students who live very close to the test center (travel time less than 5 minutes) as the baseline group. We also include our main variable of interest — non-home school status — in this specification. The results are presented in Figure 2. We find that longer travel distances are generally associated with significantly lower test scores, with the most pronounced negative impact observed among students living farthest from the test center. Moreover, the effect of testing in an unfamiliar environment is approximately equivalent to traveling an additional 25 minutes.

In sum, an unfamiliar environment — a common factor in many high-stakes assessments — is one primary driver of the observed performance penalty in our context. However, we also find that longer travel distances are negatively associated with cognitive performance, operating independently of the unfamiliar-environment effect. While this factor does not explain the observed performance penalty associated with non-home schools in our setting — where test centers are relatively concentrated — it could be more relevant in contexts where traveling to test centers requires a significant amount of time.

Figure 2: Travel Time and Academic Performance



Notes: This figure illustrates the relationship between travel time and academic performance, using students who live very close to the test center (travel time less than 5 minutes) as the baseline group. 95% confidence intervals based on high school class clusters are reported.

Lastly, a substantial body of research has examined how temporary environmental factors — such as temperature, air pollution, and pollen — can affect test performance (Marcotte, 2015; Ebenstein, Lavy and Roth, 2016; Bensnes, 2016; Graff Zivin, Hsiang and Neidell, 2018; Graff Zivin et al., 2020a; Park, 2022). However, these factors are unlikely to explain our main findings. First, the test centers in our sample are all located in the central part of the county, a relatively compact area with limited variation in environmental conditions. Second, our preferred specification includes test-room fixed effects, which account for environmental differences at a highly granular level, including room-specific characteristics such as floor level, facility conditions, and other location-specific features.

4.2 Understanding implications on disparity

In this section, we identify the most vulnerable group and explore the broader implications of our findings for educational disparities.

Heterogeneity by Student Background. We begin by examining whether low-performing students are more sensitive to unfamiliar environments than their high-performing peers. Students with higher academic ability may possess greater confidence and task focus, making them less susceptible to unfamiliar testing environments. Students are classified based on their total exam scores, with those scoring above the median defined as high achievers.[20] The results are presented in Columns (1) and (2) of Panel A, Table 4. We find that our main results are primarily driven by low achievers: their total test scores are 0.175 standard deviations lower when they take the exam in a non-home school. In contrast, the negative effect of a non-home school setting is substantially smaller for high achievers.

Furthermore, we investigate whether the performance penalty associated with the exam location differs for students from different socioeconomic backgrounds. Although students do not report their household income when registering for the exam, they are required to provide their home address in order to receive a potential offer letter. We match their home address to the average price of the neighborhood and classify students as belonging to a low socioeconomic class if the price of their neighborhood is below the median house price (i.e., ¥2500 per square meter). The results, presented in Columns (3) and (4) of Panel A in Table 4, show that students from low socioeconomic backgrounds are more strongly affected by testing

[20] Our classification may be endogenous to test center assignments, as it is based on realized test scores from the college entrance exam. However, this approach is likely to introduce only moderate measurement error. As shown in Table 1, taking the exam at a non-home institution alters a student's rank percentile by an average of just 3 points. To ensure robustness, we also adopt an alternative definition of high and low achievers by excluding students ranked in the 45th to 55th percentiles. Specifically, we define high achievers as those ranked in the 0–45th percentile and low achievers as those in the 55th–100th percentile. The results remain similar under this alternative classification.

Table 4: Heterogeneity Analysis

	Panel A					
	Total Score		Total Score		Exam Retake	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Low achievers	High achievers	Low SES	High SES	Low SES	High SES
Non-Home School	-0.175***	-0.0313**	-0.171***	-0.101***	0.0281	0.237***
	(0.0269)	(0.0152)	(0.0297)	(0.0270)	(0.0171)	(0.0213)
Statistical difference	P-value<0.001***		P-value=0.076*		P-value<0.001***	
Observations	4,135	4,379	4,070	4,459	2,733	2,882
R-squared	0.368	0.450	0.569	0.555	0.169	0.198
Test Room FEs	X	X	X	X	X	X
Highschool Class FEs	X	X	X	X	X	X
Mean of dep. var.	-0.569	0.890	0.128	0.232	0.110	0.280
	Panel B					
	College Admission		College Admission		Final College Admission	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Low achievers	High achievers	Low SES	High SES	Low SES	High SES
Non-Home School	-0.0453**	-0.00363	-0.0428***	-0.0328**	-0.0317*	0.0118
	(0.0201)	(0.00763)	(0.0146)	(0.0155)	(0.0194)	(0.0198)
Statistical difference	P-value=0.034**		P-value=0.240		P-value=0.074*	
Observations	4,135	4,379	4,070	4,459	2,733	2,882
R-squared	0.461	0.479	0.544	0.497	0.493	0.407
Test Room FEs	X	X	X	X	X	X
Highschool Class FEs	X	X	X	X	X	X
Mean of dep. var.	0.384	0.947	0.658	0.689	0.685	0.765

Notes: This table presents our additional results. $Non\text{-}Home\ School$ is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. Student socioeconomic status (SES) is classified based on whether the average housing price in their neighborhood falls above/below the county median of ¥2,500 per square meter. $Total\ Score$ represents the student's total test score, standardized by year and academic track. $Exam\ Retake$ is a binary variable indicating whether student i retook the exam in the following year. $College\ Admission$ is a binary variable indicating whether student i is admitted to any college in the current year. $Final\ College\ Admission$ is a binary variable indicating whether student i is admitted to any college either in the current year or the following year if they retook the exam. Standard errors are clustered at the high school class level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

in a non-home school.

We then examine the heterogeneity in exam retaking. While retaking the SAT boosts scores and four-year college enrollment, Goodman, Gurantz and Smith (2020) find that low-income students are 21 percentage points less likely to retake the exam potentially due to financial barriers. In their setting, addressing this gap could close up to 10% of the income-based disparity in four-year college enrollment.[21] While Table 1 shows that students assigned to non-home schools are more likely to be dissatisfied with their scores and retake the exam the following year, this average effect masks substantial heterogeneity across socioeconomic groups. Retaking the NCEE involves substantial costs, both in terms of time and financial resources. Since the exam is administered only once a year, students must spend an additional year preparing. Public high schools do not admit NCEE retakers, so these students typically enroll in private schools, which cost ¥15,000 in our sample county — where GDP per capita was approximately ¥25,000 in 2016. In addition to tuition, retakers forgo income they could have earned during that year, a burden that may be particularly significant for students from low socioeconomic backgrounds. We report results in Columns (5) and (6) of Panel A in Table 4. We find that the retaking is largely driven by students from higher socioeconomic backgrounds, who are 23.7 percentage points more likely to retake the exam when assigned to a non-home test center. In contrast, students from lower socioeconomic backgrounds are only 2.8 percentage points more likely to retake the exam, likely due to financial constraints imposed by the high costs associated with retaking.

We close here by exploring the implications of inequality on admission outcomes. Panel B of Table 4 presents the results. In Column (1), we find that low achievers are 4.53 percentage points less likely to be admitted to college if assigned to a non-home test center. However, high-achieving students are not affected in terms of college admission, likely because they are not marginal students near the admission cutoff and their test performance is less influenced by taking the exam at a non-home test center. Moreover, in Columns (3) and (4), we find that students from low socioeconomic backgrounds are more adversely affected by testing in a non-home school with respect to college admission. In Columns (5) and (6), we further account for exam retakes by constructing a new variable — final college admission — which captures

[21] Similarly, in China, retaking the NCEE also tends to improve scores and lead to better college admission outcomes on average. However, unlike the SAT — where students can submit their highest score, and retaking generally increases admission chances — the NCEE system is more uncertain. Retaking the NCEE means competing against a new cohort of students with new scores, and even with improved performance, better college admission outcomes are not guaranteed.

admission outcomes both in the current year and the following year for exam retakers. We find that when accounting for retakes, students from high socioeconomic backgrounds are much less affected by their initial test center assignment, likely because they have more resources to retake the exam if unsatisfied with their first attempt. However, for students from low socioeconomic backgrounds, the initial test center assignment still significantly influences their final college admission outcomes.

Uneven Distribution of Test Centers. In China, test centers tend to be concentrated in more advantaged high schools. To provide a more comprehensive understanding, we present a quantitative comparison of high schools that serve as test centers and those that do not in Table A5. Students from high schools that do not serve as test centers are more likely to come from disadvantaged backgrounds: they are more likely to come from low-SES families, must travel greater distances to reach test centers, exhibit significantly poorer academic performance, and are less likely to be admitted to any (elite) college. It is important to note that our main analysis focuses primarily on students from high schools designated as test centers. While this approach provides meaningful variation in test center allocation for the purpose of causal identification, it excludes students from high schools that do not serve as test centers, who may be even more vulnerable to testing in an unfamiliar environment. In fact, the average performance of students from non-test-center high schools is slightly lower than that of low-achieving students — those below the median — from test-center high schools (-0.59 vs. -0.57). Therefore, our baseline estimates likely represent a lower bound of the overall cognitive performance penalty associated with non-home test locations.

Lastly, we conduct a simple back-of-the-envelope calculation to assess the extent to which the performance gap between students from test-center high schools and those from non-test-center schools can be explained by exam location. The performance gap is 0.767 (0.182 v.s. -0.586). In our sample county, 82.9% of students from test-center high schools took the exam at a non-home school and on average experience a performance decline by 0.140. On the other hand, 100% of students from non–test-center high schools took the exam at a non-home school. Because this group lacks meaningful variation in test location, we cannot directly estimate the effect of testing away from home for them. Instead, we assume that the penalty for students from non-test-center schools is similar to that of low-achieving students in test-center schools (Table 4) as an approximate estimate, given the similarity in average NCEE performance between the two groups. Accordingly, we assume that these students, on average, experienced a performance decline of 0.175.[22] In a counterfactual scenario where all students take the exam at their home school, those from test-center high schools would score an average of 0.298, while students from non-test-center high schools would score −0.411, resulting in a performance gap of 0.709.[23] Therefore, 7.6% of the observed performance gap could be attributed to the exam location.

In our sample county, which has five test centers, approximately 17.1% of students from test-center high schools were assigned to take the exam at ahome school under fully randomized conditions. However, in counties with fewer test centers, a larger proportion of students from test-center high schools may end up taking the exam at their home schools. As a result, the performance penalty associated with the exam location could further exacerbate the realized performance gap between students from test-center and non-test-center high schools. Similarly, in a context where students are allowed to freely choose their preferred test center such as SAT, we might even expect a larger performance gap between students from test-center and non-test-center high schools.

More broadly, the concentration of test centers in high-performing schools is not unique to China. For example, in the U.S., Bulman (2015) finds that students attending high schools without an SAT test center are more likely to come from disadvantaged backgrounds — characterized by higher eligibility for free lunch, greater representation of underrepresented groups, lower academic performance, and reduced college admission rates. Recognizing this, some policymakers have, in practice, taken steps to improve equal access through better test design. Since 2000, several U.S. states have implemented mandates requiring high school juniors to take a college entrance exam (e.g., the SAT or ACT), a policy that has increased four-year college enrollment rates, particularly among students from underrepresented minority (URM) groups (Klasik, 2013; Hurwitz et al., 2015; Goodman, 2016; Hyman, 2017). When a state mandates ACT/SAT testing, it typically funds and administers the exam during school hours at students' home schools. While this approach removes logistical and financial barriers that might otherwise deter some students from taking the exam (as shown in previous studies), our findings suggest that improved educational outcomes may also be partly attributed to students performing better when taking the test at their home schools. Our findings thus contribute a new evidence-based perspective to these ongoing policy discussions.

[22] We acknowledge that our current approximate approach may not fully capture the exact penalty, and that the true magnitude could be larger, as students from non-test-center high schools also tend to travel longer distances and are more likely to come from lower-SES families.

[23] (1-0.829)*0.298+0.829*(0.298-0.140)=0.182.

5 Conclusions

Significant disparities in access to college admission exams exist among students across countries, with disadvantaged students more likely to go beyond their home schools and test in unfamiliar environments.. In this paper, we analyze administrative data from a Chinese county that randomly assigned students to test centers for their college entrance examination between 2016 and 2018. We examine whether taking the exam at a non-home school has a significant impact on student performance.

We document that students assigned to a non-home school score 0.14 standard deviations lower than those taking the exam at their home school, which in turn affects opportunities to college admission. The observed performance penalty is primarily driven by the unfamiliar environment, while longer travel distances are also associated with poorer performance. Importantly, this decline in cognitive performance disproportionately affects less-privileged groups. First, when assigned to an unfamiliar test center, lower-performing students and those from lower socioeconomic backgrounds experience greater performance drops. Second, the concentration of test centers in high-performing schools implies that less-privileged students are more likely to take exams in distant, unfamiliar locations — further widening the achievement gap between advantaged and disadvantaged groups.

While we leverage the Chinese context for identification purposes, our findings may have broader implications. Testing in unfamiliar environments are common challenges in many high-stakes contexts. Exams such as the SAT (Scholastic Assessment Test), GRE (Graduate Record Examinations), TOEFL (Test of English as a Foreign Language), IELTS (International English Language Testing System), CFA (Chartered Financial Analyst exam), and many other qualification tests often require students to travel to other cities and test in highly unfamiliar environments, imposing both financial and psychological burdens. To mitigate performance penalties associated with test location, it may be beneficial to designate more low-performing schools as test centers, if administratively feasible. Additionally, students could be given more time before the exam to familiarize themselves with the test center environment. Furthermore, technological advancements — such as the growing feasibility of at-home or online testing — may offer promising avenues for reducing these barriers. Future research could investigate whether these recent technological innovations reduce cognitive disruption in high-stakes settings and evaluate their potential role in mitigating educational inequality.

References

Beilock, Sian L, and Thomas H Carr. 2005. "When high-powered people fail: Working memory and "choking under pressure" in math." *Psychological science*, 16(2): 101–105.

Bensnes, Simon Søbstad. 2016. "You sneeze, you lose:: The impact of pollen exposure on cognitive performance during high-stakes high school exams." *Journal of Health Economics*, 49: 1–13.

Bond, Timothy N., Jillian B. Carr, Analisa Packham, and Jonathan Smith. 2022. "Hungry for Success? SNAP Timing, High-Stakes Exam Performance, and College Attendance." *American Economic Journal: Economic Policy*, 14(4): 51–79.

Borcan, Oana, Mikael Lindahl, and Andreea Mitrut. 2017. "Fighting Corruption in Education: What Works and Who Benefits?" *American Economic Journal: Economic Policy*, 9(1): 180–209.

Bulman, George. 2015. "The effect of access to college assessments on enrollment and attainment." *American Economic Journal: Applied Economics*, 7(4): 1–36.

Cai, Xiqian, Yi Lu, Jessica Pan, and Songfa Zhong. 2019. "Gender Gap under Pressure: Evidence from China's National College Entrance Examination." *The Review of Economics and Statistics*, 101(2): 249–263.

Chang, Eunsik, and Maria Padilla-Romo. 2023. "When Crime Comes to the Neighborhood: Short-Term Shocks to Student Cognition and Secondary Consequences." *Journal of Labor Economics*.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution." *American Economic Journal: Applied Economics*, 8(4): 36–65.

Frisancho, Veronica, Kala Krishna, Sergey Lychagin, and Cemile Yavas. 2016. "Better luck next time: Learning through retaking." *Journal of Economic Behavior & Organization*, 125: 120–135.

Goodman, Joshua, Oded Gurantz, and Jonathan Smith. 2020. "Take Two! SAT Retaking and College Enrollment Gaps." *American Economic Journal: Economic Policy*, 12(2): 115–58.

Goodman, Sarena. 2016. "Learning from the Test: Raising Selective College Enrollment by Providing Information." *The Review of Economics and Statistics*, 98(4): 671–684.

Graff Zivin, Joshua, Solomon M. Hsiang, and Matthew Neidell. 2018. "Temperature and Human Capital in the Short and Long Run." *Journal of the Association of Environmental and Resource Economists*, 5(1): 77–105.

Graff Zivin, Joshua, Tong Liu, Yingquan Song, Qu Tang, and Peng Zhang. 2020*a*. "The unintended impacts of agricultural fires: Human capital in China." *Journal of Development Economics*, 147: 102560.

Graff Zivin, Joshua, Yingquan Song, Qu Tang, and Peng Zhang. 2020*b*. "Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China." *Journal of Environmental Economics and Management*, 104: 102365.

Guo, Shiqi, Xinzheng Shi, and Ming-ang Zhang. 2024. "Sex Imbalance and Female Resilience: Insights from China's College Entrance Examinations." *Available at SSRN 4720789*.

Heissel, Jennifer A, and Samuel Norris. 2018. "Rise and shine: The effect of school start times on academic performance from childhood through puberty." *Journal of Human Resources*, 53(4): 957–992.

Hurwitz, Michael, Jonathan Smith, Sunny Niu, and Jessica Howell. 2015. "The Maine Question: How Is 4-Year College Enrollment Affected by Mandatory College Entrance Exams?" *Educational Evaluation and Policy Analysis*, 37(1): 138–159.

Hyman, Joshua. 2017. "ACT for All: The Effect of Mandatory College Entrance Exams on Postsecondary Attainment and Choice." *Education Finance and Policy*, 12(3): 281–311.

Jia, Ruixue, and Hongbin Li. 2021. "Just above the exam cutoff score: Elite college admission and wages in China." *Journal of Public Economics*, 196: 104371.

Kang, Le, Ziteng Lei, Yang Song, and Peng Zhang. 2024. "Gender Differences in Reactions to Failure in High-Stakes Competition: Evidence from the National College Entrance Exam Retakes." *Journal of Political Economy: Microeconomics*, 2(2): 355–397.

Klasik, Daniel. 2013. "The ACT of Enrollment: The College Enrollment Effects of State-Required College Entrance Exam Testing." *Educational Researcher*, 42(3): 151–160.

Li, Hongbin, and Xinyao Qiu. 2023. "Heuristics in Self-Evaluation: Evidence from the Centralized College Admission System in China." *The Review of Economics and Statistics*, 1–36.

Li, Hongbin, Lingsheng Meng, Kai Mu, and Shaoda Wang. 2024. "English language requirement and educational inequality: Evidence from 16 million college applicants in China." *Journal of Development Economics*, 168: 103271.

Marcotte, Dave E. 2015. "Allergy test: Seasonal allergens and performance in school." *Journal of health economics*, 40: 132–140.

Nejati, Vahid. 2023. "Chapter two - Primary principles of cognitive rehabilitation." In *Principles of Cognitive Rehabilitation.* , ed. Vahid Nejati, 59–137. Academic Press.

Park, R Jisung. 2022. "Hot Temperature and High-Stakes Performance." *Journal of Human Resources*, 57(2): 400–434.

Smith, Steven M. 1979. "Remembering in and out of context." *Journal of Experimental Psychology: Human Learning and Memory*, 5(5): 460.

Wang, Ao, Shaoda Wang, and Xiaoyang Ye. 2023. "When Information Conflicts with Obligations: the Role of Motivated Cognition." *The Economic Journal*, 133(654): 2533–2552.

6 Appendix

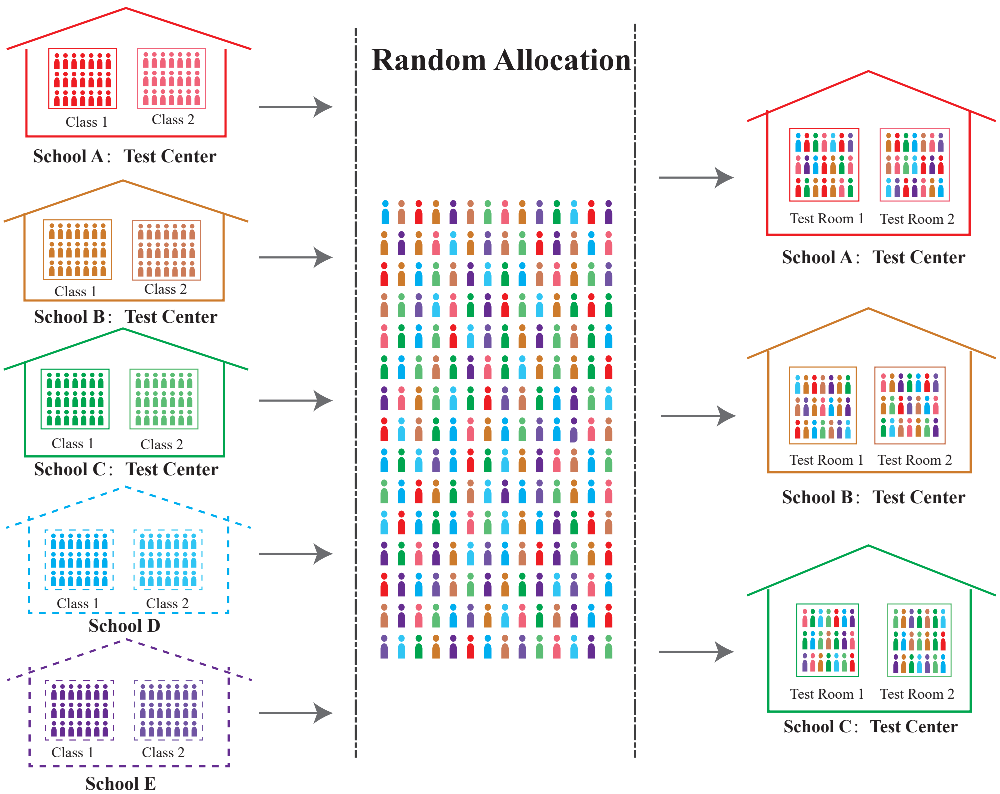
A Discussion of Major Choice

In this section, we examine whether the decline in STEM performance at a non-home school is associated with students' subsequent choices of college type or field of study (in particular, the likelihood of selecting a STEM major).

Some colleges are more technically oriented, such as the University of Science and Technology of China. We define whether a college is technically oriented based on its name. Specifically, if the name contains keywords such as "Science and Technology," "Engineering," "Medical," "Polytechnic," "Agricultural," "Pharmaceutical," "Architecture," or "Institute of Technology," we classify it as technically oriented.

In addition to selecting universities, students in China must also choose specific majors when completing their college applications. Most majors are categorized by academic track and are typically available only to either liberal arts or science students. For example, majors such as physics, engineering, and computer science are generally restricted to science-track students, while fields like Chinese literature and history are limited to those on the liberal arts track. Some interdisciplinary or broadly defined programs, such as economics or law, accept applicants from both tracks. In other words, whether a student pursues a STEM or non-STEM major in university is largely determined by their choice of the science or liberal arts track after their first year of high school (Grade 10). However, science-track students have some flexibility to pursue non-STEM majors in college. Results are reported in Panel A of Table A3. We do not find evidence that taking the exam at a non-home school reduces the likelihood of choosing a technically-oriented college or a STEM major.

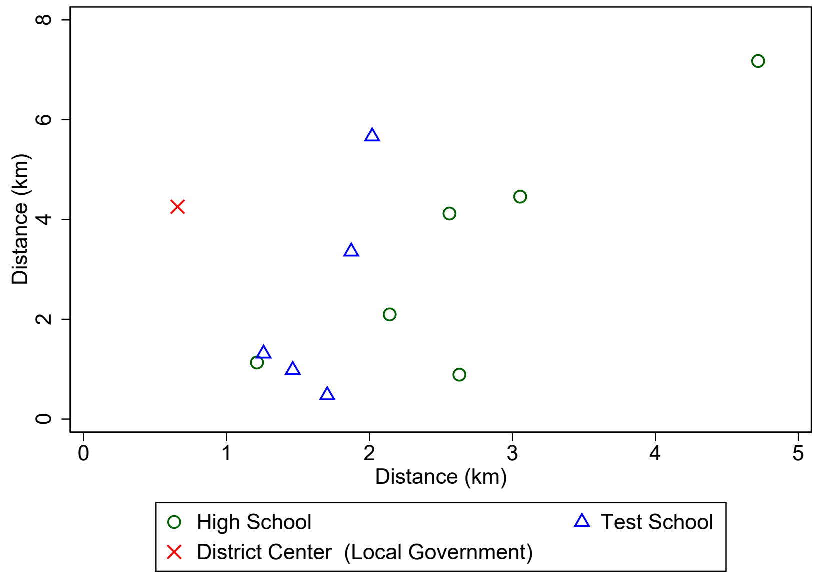
Figure A1: Random Allocation of Students to Test Centers



Notes: This figure illustrates the process of randomly allocating students to test centers. First, all students within the same county — whether from high schools that host test centers or those that do not — are pooled together. They are then randomly assigned across the available test centers, meaning that students from the same high school class may end up taking the exam in different locations.

[Back to Page 2]

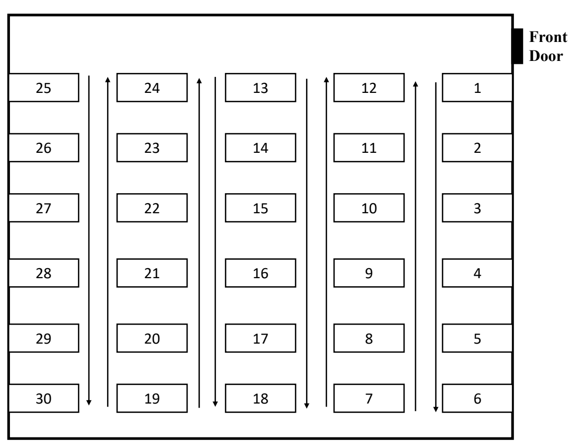
Figure A2: Relative Locations of High Schools in the Sample County



Notes: This figure depicts the relative locations of high schools and designated test centers within the sample county, since disclosure of the actual map is not permitted. Green circles represent all high schools, while blue triangles indicate the five that also serve as test centers. The red cross marks the location of the local government. The origin point is simply chosen for ease of visualization. Both axes measure distance in kilometers, illustrating the relative positions of these locations.

[Back to Page 9]

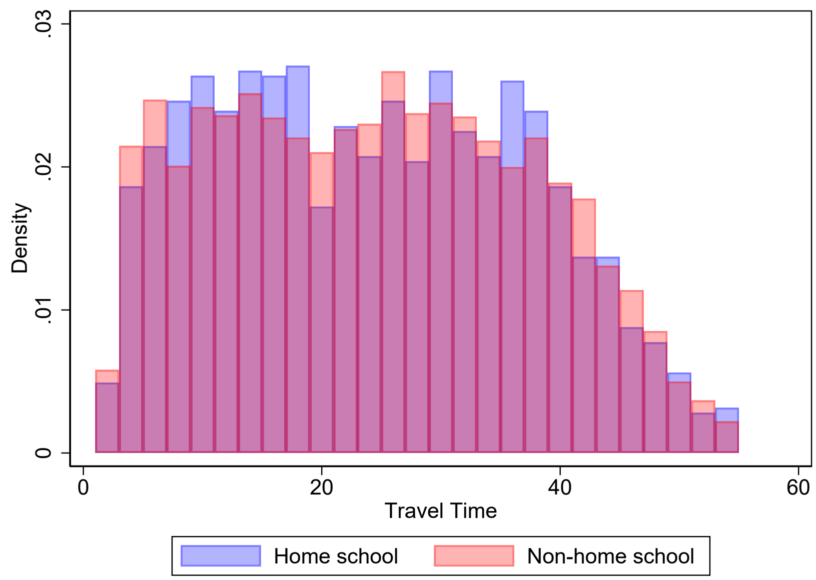
Figure A3: Floor Plan of the Test Room



Notes: The figure shows the seating arrangement of a standardized test room used in the study. The room is organized into five columns, each consisting of six rows of desks.

[Back to Page 18]

Figure A4: Distribution of Travel Time



Notes: The figure displays the distribution of travel time by test center. The x-axis represents travel time, in minutes, from the student's residential address to the test center by e-bike.

[Back to Page 20]

Table A1: Socioeconomic Characteristics of the Sampled County

	GDP (10,000 CNY)	Fiscal income (10,000 CNY)	Fiscal expenditure (10,000 CNY)	High schoolers (share of pop.)	Schooler-Teacher ratio in high schools
Year: 2016					
Sampled county	2,753,981	81,094	481,908	4.3%	14.5
Average same-province county	2,760,522	110,193	351,205	4.4%	15.4
Average Chinese county	2,235,025	167,878	344,827	4.4%	13.7
Year: 2017					
Sampled county	2,820,765	85,413	509,171	4.3%	14.5
Average same-province county	2,922,147	109,135	377,517	4.5%	15.1
Average Chinese county	2,377,611	175,838	375,689	4.5%	13.4
Year: 2018					
Sampled county	2,913,472	90,765	557,982	4.5%	14.3
Average same-province county	2,922,147	109,135	377,517	4.5%	14.7
Average Chinese county	2,583,394	184,745	409,540	4.6%	13.1

Notes: This table presents the socioeconomic conditions of our sampled county. For comparison, we also provide corresponding statistics for an average county in the same province and for an average county in China. Data source: national, provincial, and county statistical yearbooks.

[Back to Page 9]

Table A2: Balance Checks

	(1) Male (binary)	(2) Student age	(3) Urban household (binary)	(4) Class monitor (binary)	(5) Science track (binary)	(6) High house price (binary)	(7) Classmate (binary)
Non-Home School	0.0146	-0.00345	-0.00582	-2.90e-05	0.00602	-0.0219	-0.00586
	(0.0144)	(0.0124)	(0.0140)	(0.00713)	(0.0112)	(0.0143)	(0.0118)
Observations	8,535	8,535	8,535	8,535	8,535	8,535	8,535
R-squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Mean of dep. var.	0.493	17.59	0.386	0.0657	0.817	0.523	0.211

Notes: This table presents balance checks for various demographic and background variables, comparing students taking exams in home schools versus non-home schools. $Non\text{-}Home\ School$ is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. Robust standard errors are reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

[Back to Page 12]

Table A3: Additional Results

	Panel A: Additional Outcomes		
	(1)	(2)	(3)
VARIABLES	Raw Total Score	Technically-Oriented College	Stem Major
Non-Home School	-10.13***	0.0105	-0.0161
	(1.390)	(0.0139)	(0.0168)
Observations	8,535	7,848	5,211
R-squared	0.554	0.182	0.139
Test Room FEs	X	X	X
Highschool Class FEs	X	X	X
Cluster level	Highschool Class	Highschool Class	Highschool Class
Sample	Main Sample	College-Admitted Students	College-Admitted Students from the Science Track

	Panel B: Additional Specifications		
	(1)	(2)	(3)
VARIABLES		Total Score	
Non-Home School	-0.141***	-0.141***	-0.153***
	(0.0214)	(0.0192)	(0.00440)
Total Score (Last Year)			0.975***
			(0.00902)
Observations	8,535	8,535	1,065
R-squared	0.512	0.512	0.996
Test Room FEs	X	X	.
Highschool Class FEs	X	X	.
School-Track-Year FEs	.	.	X
Cluster level	High school * Year * Track	Test Room	High school * Year * Track
Sample	Main Sample	Main Sample	Repeat Exam Takers

Notes: This table presents our additional results. $Non\text{-}Home\ School$ is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. In Panel A, Column (1) uses raw total test scores as the outcome variable, while Columns (2) and (3) use college admission to technically-oriented programs and whether the major is in STEM, respectively. Standard errors are clustered at the high school class level. In Panel B, $Total\ Score$ represents the student's total test score, standardized by year and academic track. $TotalScore(LastYear)$ refers to an individual's standardized total test score from the previous year. Standard errors are clustered at the school-year-track level in Columns (1) and (3), and at the test room level in Column (2). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

[Back to Page 13]

Table A4: Additional Results by Subgroups

	Panel A: Results by Year and Gender				
	(1)	(2)	(3)	(4)	(5)
	Total Score			Total Score	
Sample	2016	2017	2018	Male	Female
Non-Home School	-0.134***	-0.125***	-0.164***	-0.152***	-0.129***
	(0.0368)	(0.0340)	(0.0260)	(0.0272)	(0.0263)
Observations	2,762	2,857	2,916	4,203	4,328
R-squared	0.507	0.508	0.520	0.613	0.508
Test Room FEs	X	X	X	X	X
Highschool Class FEs	X	X	X	X	X

	Panel B: Results by School				
	(1)	(2)	(3)	(4)	(5)
	Total Score				
Sample	School A	School B	School C	School D	School E
Non-Home School	-0.108***	-0.127***	-0.146***	-0.107***	-0.202***
	(0.0310)	(0.0429)	(0.0459)	(0.0380)	(0.0395)
Observations	2,283	1,746	1,661	1,208	1,637
R-squared	0.428	0.424	0.447	0.390	0.484
Highschool Class FEs	X	X	X	X	X
Mean of "Non-Home School"	0.823	0.848	0.827	0.801	0.839

Notes: This table presents our additional results by subgroups. $Non\text{-}Home\ School$ is a binary variable indicating whether student i is assigned to a test center outside their home school to take the exam. $Total\ Score$ represents the student's total test score, standardized by year and academic track. Standard errors are clustered at the high school class level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

[Back to Page 16]

Table A5: Summary Statistics for High Schools

	High schools as test centers	High schools not as test centers	Mean difference
Panel A: Student characteristics			
Male	0.493	0.454	0.039***
	(0.500)	(0.498)	(0.000)
Han ethnicity	0.997	0.998	-0.001
	(0.053)	(0.047)	(0.552)
Age	17.588	17.594	-0.006
	(0.436)	(0.414)	(0.010)
Urban household	0.386	0.324	0.061***
	(0.487)	(0.468)	(0.000)
High-housing-price community	0.523	0.485	0.038***
	(0.500)	(0.500)	(0.000)
Science track	0.817	0.501	0.316***
	(0.387)	(0.500)	(0.000)
Distance to test centers	6.144	6.308	-0.163**
	(3.244)	(3.225)	(0.025)
Panel B: College entrance exam scores (standardized by year $\times$ track)			
Total	0.182	-0.586	0.767***
	(0.939)	(1.047)	(0.000)
Chinese	0.137	-0.437	0.574***
	(0.972)	(1.002)	(0.000)
Math	0.165	-0.534	0.698***
	(0.945)	(1.035)	(0.000)
English	0.136	-0.447	0.583***
	(0.974)	(0.997)	(0.000)
Comprehensive (STEM)	0.181	-0.826	1.007***
	(0.955)	(0.950)	(0.000)
Comprehensive (Non-STEM)	0.082	-0.222	0.304***
	(0.956)	(1.021)	(0.000)
Panel C: College admission			
Any college	0.674	0.248	0.426***
	(0.469)	(0.432)	(0.000)
Elite college	0.016	0.002	0.014***
	(0.124)	(0.043)	(0.000)

Notes: This table presents summary statistics for students attending high schools that do and do not serve as college entrance exam testing centers. Each cell reports the mean, with the standard deviation in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

[Back to Page 25]