The End of One Exam: Centralized Assignment Reform, Student Choice, and School Quality*

Rui Zeng† — *Job Market Paper*
Victor Delgado‡
Jintao Sun§

Last Updated: October 25, 2025

Click here for the most updated version.

Abstract

This paper studies how centralized assignment rules shape student placement and school quality. We examine Mexico's 2025 reform, which replaced exam-based assignment for non-elite schools with a two-list mechanism. Using administrative applicant data and standardized test records, we estimate school value-added as a measure of quality and build a structural model of school demand and supply. On the demand side, we model students' school choices and estimate preference parameters under the Stability assumption. On the supply side, schools determine their quality level by trading off enrollment incentives and direct returns to quality against their associated costs. Estimates show students prefer high quality schools. Non-elite schools improve quality only in response to enrollment incentives. In contrast, elite schools respond to both enrollment incentives and direct benefits from providing higher quality. Counterfactuals show that the new system shifts high achievers to non-elite schools; however, it also increases the number of unassigned students. The new system changes school quality, with value-added falling at elite schools and rising at non-elite schools. Yet elite schools continue outperforming non-elite schools. Expanding non-elite capacity reduces unassigned students without altering these effects. This paper highlights both the equity gains and the capacity challenges of the reform.

Keywords: *School Choice, Market Design, Centralized Matching Market, Student Preferences, School Competition, Mexico*

JEL Codes: D12, D47, I21, I24, J18, L31

*We are deeply indebted to YingHua He for his invaluable guidance and encouragement throughout this project. We thank Rossella Calvi, Flávio Cunha, Ying Fan, and Isabelle Perrigne for their valuable comments. We also thank Rafael de Hoyos and Xaber A.C. for providing access to the data.

†Department of Economics, Rice University; Email: rz32@rice.edu.

‡Department of Economics, Rice University; Email: victor.delgado@rice.edu.

§School of Applied Economics, Renmin University of China; Email: jintaosun@ruc.edu.cn.

1 Introduction

Over the past two decades, many public school systems have shifted from ad hoc applications to a centralized school choice system, combining common application platforms with algorithmic assignment. While other market-oriented reforms, such as vouchers and charter expansion, remain contested, these coordinated assignment systems have quietly become a standard practice, restructuring choices without building more schools. Their use now spans major U.S. districts (e.g., Boston, New York City, New Orleans, Denver) and international settings (e.g., college admissions in China, France's Affelnet and Parcoursup systems, and Chile's nationwide system), signaling a broad and worldwide adoption of mechanism-based school assignment (Neilson, 2024). A centralized school choice appeals to policymakers because it simplifies admissions with a single ranked application and transparent rules, and it loosens the residence–school link, expanding access and improving the fitness between schools and students. By widening families' options, these systems are also expected to encourage schools to improve quality in response to competitive pressure.

However, the evidence is mixed on both students and school margins. On the student side, evidence on how school choice affects enrollment student sorting and segregation is mixed. Centralized admissions can weaken the residence–school link and open access across neighborhoods (e.g., Hastings et al., 2006; Abdulkadiroğlu et al., 2009), but priority rules, such as sibling, walk-zone, or test-score based admissions, can reintroduce stratification. Indeed, studies find persistent or even intensified socioeconomic sorting under these designs (e.g., Cullen et al., 2006; Urquiola, 2005; Pathak and Sönmez, 2013) . On the school side, evidence on whether competition improves school quality is similarly mixed. Greater competition between districts has been associated with higher student academic outcomes (Hoxby, 2000; Card et al., 2010). Studies on charter school entry suggest limited or null competitive responses in nearby public schools (Zimmer and Buddin, 2009; Imberman, 2011; Ni, 2009).

There is little direct evidence on how the design of centralized assignment rules (e.g., priorities, tie-breakers, or algorithm choice) affects school quality. These designs matter because they shape the school's competitive environment. For example, in a centralized system with exam-based priority assignment, high-scoring students are typically allocated to top schools, while students with lower scores are assigned to lower-quality schools. However, when the same system switches to a pure lottery assignment, all applicants have an equal chance to be considered at any school, including low-scoring students at high-quality schools. If students prefer higher quality, this mechanism exposes low-quality schools to more intense competition. As a result, they have stronger incentives to improve and attract students, increasing the perceived payoff from investing in quality.

Motivated by mixed findings and limited evidence on admission rules, this paper presents new evidence on how rule design in a centralized school choice system affects student allocation and school quality responses. Specifically, this paper studies a reform that eliminates test-score priorities and introduces a two-list mechanism, allowing students to submit separate rank-order lists for elite and non-elite schools, in the admission system organized by the Comisión Metropolitana de Instituciones Públicas de Educación Media Superior (COMIPEMS). This centralized assignment allocates public high school seats across Mexico City's metropolitan area. This paper (i) quantifies how changes to assignment rules reconfigure who receives offers and where, and how this re-mapping affects student composition across schools; and (ii) documents how schools adjust quality in response to the altered competitive environment. We find that under the new admission rule, a meaningful share of high achievers shift to non-elite schools, while overall filled enrollment declines. Schools also adjust their qualities: value-added falls at elite schools and rises at non-elite schools, although elite schools continue to outperform on average, consistent with prestige effects. Overall, the reform brings equity gains but creates capacity pressures, emphasizing the need for future assignment designs to account for additional factors such as capacity constraints.

Until the academic year of 2025, COMIPEMS centrally allocates public upper-secondary seats using a single entrance exam and a rank-ordered list (ROL) of preferred schools.The market is large and heterogeneous, spanning three major educational tracks (general academic, technical, and vocational) and a mix of elite and non-elite schools, with substantial differences in selectivity and program offerings. Students participating in this system vary widely in their socioeconomic backgrounds and prior academic achievements. These features generate meaningful variations in school quality and in student sorting. To gain admission, applicants submit a ROL of high schools. Before 2025, all applicants took a unified placement exam, and the resulting score was the sole priority across all schools. Seats are allocated via a Serial Dictatorship (SD) mechanism, where students are ordered by score and assigned to their highest-ranked program with available capacity. Starting in summer 2025, the admission scheme drops score-based priority for non-elite schools and allows two ROLs, one for elite schools and one for non-elite schools. For non-elite programs, if applications do not exceed capacity, all listed applicants are admitted. If oversubscribed, seats are allocated through a gender-balanced random draw, ensuring a 50% female intake. Elite programs remain unchanged, continuing to rank applicants based on exam scores and allocate seats in order of score.

We assemble rich administrative microdata from COMIPEMS, encompassing the universe of registered applicants (2019 cohort), their ROLs, placement test scores, and final assignments. We link these records to public high school administrative files, which contain school attributes. Finally, we use standardized test archives from the Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) to construct school-level value-added metrics, which serve as our measure of school quality. We document that under the SD mechanism, higher value-added schools generally admit stronger students at higher cutoffs.[1] They also tend to enroll a less dispersed set of students in terms of prior achievement.

To examine the 2025 policy, we develop a structural model where students' school demand on one side, and schools choose value-added on the other, both embedded in the assignment mechanism. We first estimate students' preferences over high schools using a two-sided matching framework. In particular, we adopt the assumption that the observed matching between students and high school programs is stable in equilibrium (Fack et al., 2019; Artemov et al., 2023). The stability-based approach has been shown to generate more precise preference estimates than other assumptions, such as the truth-telling assumption. We allow students to value schools based on distance, student demographics, school attributes, and value-added (quality). On the supply side, public schools choose value-added to maximize an objective that equals benefits minus costs. Benefits have two components: (i) an enrollment-driven term that rises with the number of students attracted at a given quality, and (ii) a direct, non-demand component that captures the intrinsic value of providing higher quality, such as prestige. This intrinsic component differs between elite and non-elite schools. The cost arises from the marginal expense of enrolling additional students and from the resources required to maintain quality, such as hiring high-quality teachers. This formulation implies an equation deriving the optimal choice of value-added: schools set quality at the "perfect competition" level, adjusted downward by a markdown that depends on the elasticity of demand they face (Neilson, 2021; Gilraine et al., 2023).

The model faces several identification and estimation challenges. First, the large choice data create a substantial computational burden, which we address by leveraging deep learning techniques that allow GPU acceleration and efficient handling of high-dimensional choice sets. Second, iden-

[1] A cutoff score is the minimum exam score among admitted students in a school.

tification is complicated because schools' value-added is endogenous to unobserved demand incentives. To address this identification challenge, we use competitors' classroom computer use as an instrument, which shifts a school's value-added through competitive pressure and yet is plausibly excluded from that school's own unobserved demand shock. Third, the estimation requires valid demand shifters to separate preference heterogeneity from endogenous supply responses. In particular, we need variables that influence students' school choices without simultaneously altering the unobserved costs of schools, so that the resulting exogenous variations can be used to identify school decisions. To this end, we construct shifters based on the weighted number of nearby students, the availability of public transportation, the presence of nearby schools, and the previous year's admission cutoffs.

Our preference estimates show that student demand falls sharply with distance and with higher socioeconomic status, and rises with prior achievement. Elite schools deliver higher student mean utility than all other school types. Among non-elite schools, mean utility is much more dispersed, indicating a wider heterogeneity in perceived quality. Moreover, we identify a positive and significant effect of school quality on mean value-added, confirming that higher value-added schools are associated with higher average utility and stronger revealed preference for that school. On the school side, estimates show that schools value enrollment, which responds to quality. Non-elite schools have negative net direct returns to quality, implying little incentive to improve absent enrollment pressure. Elite schools, in contrast, receive positive net direct returns from higher quality, indicating benefits beyond attracting more students, such as prestige and reputation.

Counterfactual analysis indicates that the 2025 assignment rule reshapes admissions and school quality. Eliminating the exam-based assignment rule for non-elite schools shifts a meaningful share of high achievers toward non-elite schools. The gender-balanced lottery raises female representation in non-elite schools. At the same time, the new assignment rule leaves more students without placement for two reasons. First, under the two-list system, all applicants submit a ROL only for non-elite schools, so the same group of students now compete only for non-elite seats instead of the combined pool of elite and non-elite schools. This narrower competition results in more students being left unassigned. Second, some students receive both elite and non-elite offers but accept only one, and without a second round reassignment, the declined seats remain vacant. The decline is substantial at elite schools, as many students with dual offers choose non-elite schools which improve their quality under the new equilibrium.

School quality adjusts in response: non-elite schools raise value-added as small improvements attract demand. Under the two-list system, small improvements in quality now have a more direct influence on which students apply and enroll. The assignment mechanism amplifies the payoff to quality improvements by making enrollment more responsive to perceived school quality. Elite schools reduce quality as demand decreases and smaller enrollments lower the returns to maintaining high value-added. However, they continue to outperform non-elite schools because non-enrollment incentives sustain higher intrinsic quality. Lastly, by running counterfactuals on the 2025 reform with expanded seats, we find that expanding non-elite capacity reduces the number of unassigned students while leaving the student assignment pattern and quality patterns unchanged.

To sum up, our study highlights both the equity gains and the capacity challenges of the reform, suggesting that assignment rule changes should be paired with complementary policies, such as capacity expansion, supplemental rounds, and the construction of new schools, to ensure the reform's success.

Related Literature This paper relates to and contributes to several strands of literature.

First, this paper contributes to the extensive literature that develops and estimates empirical models of school choice and competition. Building on earlier theory (e.g., Epple and Romano, 1998; McMillan, 2004), recent work has employed structural discrete choice models to investigate student preferences and school incentives under specific school choice programs. In the case of vouchers, studies such as Figlio and Hart (2014) and Figlio et al. (2023) document how the introduction of voucher programs alters both student sorting and public school performance. Neilson (2021) studies voucher-induced competition in Chile, developing a structural model where schools set quality in response to household preferences and voucher policy design. Regarding charters, Epple et al. (2021) model charters' endogenous choice of practices, including curriculum. Gilraine et al. (2023) structurally estimate that competition raises public value-added, especially when entrants use traditional, skills-focused curricula.

Our paper focuses on centralized assignment policies and their consequences for school incentives and quality. Prior work has examined how these mechanisms shape access and choice behavior (e.g., Abdulkadiroğlu et al., 2005, 2020; Hastings et al., 2006) but this literature primarily focuses on preferences and strategic behavior rather than the downstream effects of assignment rules on schools themselves. In contrast, we develop a structural framework that links assignment rules to both student demand and schools' value-added choices, allowing us to quantify how different mechanisms shape equilibrium quality.

Second, the paper also relates to the literature on estimating student preferences over schools in a centralized assignment system. While deferred acceptance mechanisms are strategy-proof in theory (Abdulkadiroğlu and Sönmez, 2003), prior research shows that assuming students' submitted ROLs truthfully reveal their preferences is restrictive (e.g., Haeringer and Klijn, 2009; Fack et al., 2019; Artemov et al., 2023). In practice, ROLs are often constrained in length, and students may strategically omit unlikely options, so treating them as exact preference orderings can lead to biased estimates. To address this, Fack et al. (2019) propose a stability-based approach that infers preferences from the observed stable matching, and subsequent studies across various matching contexts adopt this framework (e.g., Akyol and Krishna, 2017; Combe et al., 2022; Ainsworth et al., 2023; Bobba et al., 2021, 2023), showing it yields more precise and credible preference estimates than truth-telling assumptions. Our paper builds on the stability-based approach of Fack et al. (2019) to generate credible preference estimates, which in turn shape our supply-side estimates by linking demand to schools' value-added choices. While most of the existing literature focuses on demand estimation alone, our framework explicitly connects preferences and assignment outcomes to the incentives schools face in equilibrium.

Third, this paper relates to the literature on school modeling (not only price-seeking). A large body of work has documented how competitive pressures influence school performance (e.g., Hoxby, 2000; Figlio and Hart, 2014). Moreover, a growing strand of research models school behavior in an empirical IO framework, often incorporating both public and private providers. Allende (2019) and Neilson (2021) structurally estimate value-added competition across mixed systems of public and private voucher schools. In contrast, Bau (2022) shows that match quality matters for learning, while profit-maximizing private schools tend to tilt their quality toward wealthier students, thereby increasing inequality and reducing overall welfare. Our setting focuses on public schools with free tuition. Following the theoretical foundation in McMillan (2004), we follow the approach of Gilraine et al. (2023) and model schools as maximizing enrollment net of costs, treating value-added choice as the component responsive to competitive incentives.

Organization of the Paper The remainder of this paper is organized as follows. The next section introduces institutional features, such as COMIPEMS and assignment rules prior to 2025, as well as the new assignment scheme introduced in 2025. Section 3 provides detailed information on data sources in our analysis and summary statistics. Section 4 outlines our estimates of value-added and highlights key data patterns. We then present our student demand and school supply model in Section 5. Section 6 outlines identification and estimation. In Section 7, we present the estimated model parameters. Section 8 uses the model estimates to counterfactually evaluate the effects of the policy scheme introduced in 2025. Section 9 concludes.

2 Institutional Background

In this section, we introduce the high school admission system in Mexico City. We first describe the previous admission system, which is operated by the Metropolitan Commission of Public High School Institutions (COMIPEMS, by its acronym in Spanish).[2] We then discuss the 2025 admission reform that eliminates exam-based assignment for non-elite schools and introduces a two-list mechanism, focusing in particular on its potential consequences on student admission as well as school quality change.

2.1 COMIPEMS and Assignment Rules Before 2025

From 1996 until 2024, COMIPEMS organizes the admission process for students seeking to enter public higher secondary education institutions in the Metropolitan Area of Mexico City. Through this mechanism, applicants compete for seats across nine public educational subsystems, covering 38 municipalities (including the 16 municipalities of Mexico City and 22 municipalities in the state of Mexico) under a single examination and allocation system. Importantly, COMIPEMS operates in the largest educational market in Mexico nationwide, with more than 300,000 applicants registered and an annual admission rate of around 80%. The admitted students account for approximately three-quarters of all high school enrollment in the metropolitan area (Delgado et al., 2025).

Upper secondary education in Mexico is organized into three main tracks, or educational pathways. The *Bachillerato General* (general high school) track offers an academic curriculum designed to prepare students for university studies. The *Bachillerato Tecnológico* (technological high school) track combines general education with technical or professional training, granting both

[2] Throughout the paper, we use COMIPEMS to refer broadly to both the commission and the admission system it administers. The term is also used adjectivally (e.g., COMIPEMS system, COMIPEMS area, COMIPEMS exam) to describe features associated with the centralized high school admission process in Mexico City prior to 2025.

a high school diploma and a technical certificate. Finally, the *Profesional Técnico-Bachiller* (professional technical baccalaureate) track is primarily vocational, emphasizing labor market preparation through industry relevant skills while still granting an upper secondary credential.[3]

These tracks are delivered through subsystems, which are administrative units of upper secondary education. Subsystems differ by their governing authority (federal or state), their curricular orientation, or their affiliation with higher education institutions. In the literature, these subsystems are often classified into elite and non-elite. Elite subsystems include high schools directly affiliated with either the *Universidad Nacional Autónoma de México* (UNAM) or the *Instituto Politécnico Nacional* (IPN), such as UNAM's *Escuelas Nacionales Preparatorias* (ENP) and IPN's *Centros de Estudios Científicos y Tecnológicos* (CECyT). These high schools admit only the highest-scoring applicants based on the COMIPEMS entrance exam. Importantly, UNAM-affiliated high schools offer pathways to guaranteed admission into UNAM for high-performing students through the *pase reglamentado* system. In contrast, CECyT does not provide a similar guaranteed admission to IPN higher education programs; however, they are widely recognized for their rigorous technical and scientific preparation, which effectively equips students to excel in the IPN university entrance exam.

Non-elite subsystems encompass a variety of upper secondary education institutions, including the *Colegio de Bachilleres*, the *Colegio Nacional de Educación Profesional Técnica* (CONALEP), and the *Dirección General de Educación Tecnológica Industrial* (DGETI), among others. In the State of Mexico, subsystems are grouped into a single category under the Secretariat of Education of the State of Mexico (SE). Non-elite subsystems also differ in their educational objectives, offering academic, technological, and vocational programs.

In some cases, subsystems are composed of different educational institutions. For example, the DGETI includes the *Centros de Estudios Tecnológicos Industriales y de Servicios* (CETIS) and the *Centros de Bachillerato Tecnológico Industrial y de Servicios* (CBTIS), both of which combine a general high school curriculum with technological training, though they differ in the specific programs they offer. Each institution operates across multiple campuses (*planteles*), where students apply directly and, in the case of technical programs, choose a specific career track.

Table 1 summarizes the structure of the upper secondary education system participating in the joint admission process organized by COMIPEMS in 2019. Column (1) lists the participating *sub-*

[3] Vocational schools also provide pathways to higher education.

systems. Column (2) indicates whether a subsystem is classified as *elite*, a designation that applies only to UNAM and IPN, which admit the highest-scoring applicants into their selective preparatory and technological schools. Column (3) identifies the main educational institutions within each subsystem, such as CETIS and CBTIS. Column (4) reports the *track* associated with each institution, distinguishing between general academic (G), technological (T), and vocational (V). Column (5) shows the number of school sites under each institution. Finally, column (6) reports the total number of options available to applicants. An option is defined as a subsystem–school site–program combination. The number of options can exceed the number of school sites when a single site offers multiple programs, particularly technical-vocational ones. For example, CONALEP operates 28 school sites but provides 92 options, reflecting the breadth of vocational offerings available within each site.

As shown in Table 1, the size and composition of the subsystems participating in the joint admissions process considerably vary. The SE operates by far the largest number of school sites in COMIPEMS, making it the primary provider of high school options in the metropolitan area. In contrast, the elite subsystems are relatively small in terms of school sites, with only 30 school sites in total, yet they play a disproportionate role in shaping the stratification of access to high-quality education.

The high school application process organized by COMIPEMS followed a specific timeline. In late January of each year, students obtain free information about the process from national newspapers or the COMIPEMS website. Applicants also have access to a detailed list of all available educational programs, specified down to the subsystem–school site–option level, including curriculum contents and admission capacities (i.e., number of seats). From late February to early March, students complete the registration process by filling out a registration form, completing a demographic survey, and submitting a rank-ordered list (ROL) of up to 20 options.

In June, students take a standardized entrance exam administered by the National Evaluation Center (CENEVAL). The exam consists of 128 multiple-choice questions worth one point each, covering ten subjects.[4] In July, high schools reported their capacity to CENEVAL, and the merit-based Serial Dictatorship (SD) mechanism is used to match students with schools.[5] The algorithm

[4]The ten subjects include Verbal Skill, Spanish, History, Geography, Civic and Ethical Training, Mathematical Skill, Math, Physics, Chemistry, and Biology. In 2019, Verbal Skill and Mathematical Skill take up 16 points each, and all the other subjects are 12 points each, while holding a total score of 128 points.

[5]A special case of the celebrated student-proposing Deferred Acceptance (DA) mechanism (Gale and Shapley, 1962) as all the programs rank students in the same way.

orders all students based on their entrance exam scores in descending order. It then assigns each student to their highest-ranked school with available seats, beginning with the highest-scoring student and moving down the list (See Appendix A.1 for more details on the mechanism).

To be assigned to any school, students need to obtain their middle school certificate, and to be assigned to any elite option, it is necessary to obtain a minimum GPA of 7.0 in middle school. CENEVAL announces the placement test scores as well as assignment results in late July. Following this main computerized assignment, there is also a supplementary round to match part of the unassigned students (e.g., who later obtain the middle school certificate) to high school programs with remaining unfilled capacities, more in a one-on-one negotiating manner. The final assignment outcomes are announced in late July or early August (COMIPEMS, 2023).

Although students submit their ROLs and are admitted at the option level, in our analysis, we aggregate options to the school site level. The main purpose of the paper is to model school decisions and capture competition across schools rather than within-school programs. In the rest of the paper, we use "school" as the school site.

Table 1: Number of Campuses & Education Options within COMIPEMS in 2019

Subsystem (1)	Elite Status (2)	Institution (3)	Track (4)	# School sites (5)	# Options (6)
COLBACH	No	COLEGIO DE BACHILLERES	G	20	20
CONALEP	No	CONALEP CDMX	V	28	92
DGB	No	CEB	G	2	2
DGETI	No	CBTIS/CETIS	T	51	51
DGETAyCM	No	CBTA	T	6	6
		CMM	T	1	1
SE	No	CBT	T	37	110
		CECYTEM	T	26	59
		COBAEM	G	20	20
		CONALEP EM	V	29	97
		P.O.	G	181	181
		TELEBACHILLERATO	G	42	42
UAEM	No	UAEM	G	1	1
IPN	Yes	CECyT	T	16	16
UNAM	Yes	CCH/ENP	G	14	14
Sum		......		474	712

Notes: This table reports the number of school sites and education options offered by each subsystem participating in the COMIPEMS admissions process in 2019. Subsystems are listed in column (1), with column (2) indicating whether they are classified as elite. Column (3) provides the name of the institution(s) within each subsystem. Column (4) shows the educational track: general academic (G), vocational (V), or technological (T). Columns (5) and (6) report, respectively, the number of school sites and the total number of education options available. The bottom row presents the overall totals (474 school sites and 712 options). Note that the number of options can exceed the number of campuses, as a single site may offer multiple programs (e.g., CONALEP has 28 sites but 92 options).

2.2 The 2025 Admission Reform

Beginning in 2025, the long-standing COMIPEMS exam is replaced by a new admission system entitled *Mi derecho, mi lugar*, coordinated under the Espacio de Coordinación de Educación Media Superior (ECOEMS).[6] The reform was motivated by concerns over exclusion generated by the centralized exam and the desire to guarantee that all students completing lower secondary education could access upper secondary schooling.

The new system separates the admission process for elite and non-elite schools. All non-elite institutions now admit students directly without the use of an entrance exam. By contrast, two elite subsystems, UNAM and IPN, retain exam-based admissions. Students can submit two ROLs, one with up to ten educational options from institutions with direct access (which do not require an entrance exam), and one with up to five from IPN and five from UNAM, both of which require an exam to assign applicants.

The assignment procedure is separated into parts. First, for applicants who submit a list of educational options from institutions that do not require an entrance exam, they adopt a mechanism similar to the Boston mechanism (BM). The assignment of educational spaces is based on the order of their listed preferences and the number of available seats in each registered option. The assignment criterion is to prioritize the preferences indicated in their list of educational options. If the applicant's chosen educational option is a low-demand option (i.e., it has enough places to accept all its applicants), he/she is assigned directly. If it is a high-demand option (i.e., it does not have enough places for all applicants), a lottery is conducted, with a gender quota to guarantee at least 50% of the places for female applicants. Suppose the applicant is not assigned to any of the options in his/her list of educational options. In that case, he/she is offered educational alternatives that are nearby or similar to the first option listed on their registration form. See Appendix A.2 for more detailed information on the assignment procedure.

Note that the mechanism here is not exactly the same as BM. In the school choice literature, the standard BM is typically defined as an immediate-acceptance procedure. In this mechanism, oversubscribed schools admit applicants according to predetermined priority rules (e.g., exam scores, neighborhood zones), with lotteries serving only as tie-breaking devices. By contrast, the ECOEMS version of BM does not incorporate priority rules. When an option is oversubscribed, all

[6] Throughout the paper, we refer to this new admission framework and its associated features collectively as ECOEMS.

applicants are treated symmetrically and seats are allocated entirely by lottery, subject to a gender quota that guarantees at least half of the places for female students when necessary. Thus, the ECOEMS design can be described as a BM without priorities, in which the lottery itself constitutes the assignment rule rather than a secondary tie-breaking instrument.

For students who submit a list of elite options, the SD mechanism is applied based on the number of correct answers on the exam, the order of the educational preferences, and the number of available seats in each educational option. Note that this assignment process only applies to applicants who take the entrance exam and obtain a minimum secondary school GPA of 7.0.

It is important to mention that while eliminating exam-based entry appears to broaden access, it might also alter incentives for both students and schools. On the student side, because admission is now determined by lottery rather than by exam scores, more applicants might perceive high-quality non-elite schools as attainable, generating congested demand: competition intensifies, and some students may be crowded out by lottery outcomes. Peer composition also shifts as lotteries admit more mixed cohorts. On the school side, since all applicants have a chance to be considered at any school, including low-scoring students at high-quality schools, this exposes low-quality schools to a more intensified competition environment and potentially increases the perceived payoff of improving quality. Together, these dynamics suggest that the reform can reshape both student choices and school behavior by intensifying competition.

3 Data and Summary Statistics

The data used for this paper come from several sources. We primarily leverage the administrative data of application and admission records compiled by the COMIPEMS system. We obtain student-level application and assignment data for the academic years from 1996 to 2019. This dataset includes detailed information on students' ROL, their exam scores, which are essential for the centralized assignment process, and their final admission outcomes. Furthermore, the dataset is enriched by a pre-examination survey that all students are mandated to complete. This survey offers comprehensive insights into the students' backgrounds, including demographic information such as gender, whether the applicant is a Mexico City Resident, categorical monthly income, parental education, and home address. It provides data on students' academic performance in middle school, including the name and universal identifier of the middle school attended, whether the middle school is public, whether the middle school is general academic, and the students' 9th Grade Point Averages (GPAs). It also contains detailed information on the schools/options participating in COMIPEMS. More specifically, we observe the school location, track, and elite status. Besides school information, we also obtain the options offered by each school.

In addition to the applicant data, we obtain the "Formato 911" school census data that is collected by the Secretariat of Public Education (Secretaría de Educación Pública, SEP). This comprehensive school survey is part of the annual educational statistical collection process, aimed at gathering detailed information on schools throughout the country. It is collected twice, at the beginning and at the end of every school year.

The "Formato 911" school provides information for all public and private schools in Mexico, encompassing all levels of educational institutions from preschools to higher education. The "Formato 911" collects data on various aspects of educational institutions, including student enrollment numbers, teacher counts, administrative staff, infrastructure details, and other critical educational resources and services. Moreover, the school's census data contains school identifiers that allow us to match with COMIPEMS data.

We use the school census data for the following purposes. First, we can trace back various attributes of middle schools from which the registered students graduated, including regime (public and private) and track (academic or non-academic). Second, we can obtain features of public high schools that participate in COMIPEMS. These features mainly include the student-teacher ratio, gender composition of students, gender and educational composition of teachers, and some infrastructure details. Third, as a module under Formato 911, the Information Technologies Annex collects detailed information on the digital infrastructure of schools. The annex records the availability of hardware such as desktop computers, laptops, tablets, and projectors; measures of connectivity, including internet access and bandwidth; and the existence of dedicated computer laboratories or ICT classrooms.

To obtain the school quality measure, we estimate school value-added using data from the National Assessment of Academic Achievement in Schools (ENLACE, by its Spanish acronym). ENLACE is a nationwide standardized test administered annually from 2006 to 2013 in Mexico. The exam covered multiple grade levels in primary, secondary, and high school, with the goal of providing a comparable measure of student achievement across schools and regions. For our purpose, the most relevant cohort is high school students in their final year (Grade 12). The test consists of subject-specific sections in mathematics and language, complemented by contextual questionnaires that collect detailed background information on students, families, and past academic records. The dataset reports student scores standardized as z-scores, using national means (500) and standard deviation (100) as the reference distribution (de Hoyos et al., 2018).[7] For example, a z-score of 0.5 corresponds to an ENLACE test score of 550, half a standard deviation above the national mean. Importantly, participation is effectively universal for enrolled students in the tested grades, which minimizes concerns about selective take-up. Together, these features make ENLACE a valuable dataset for analyzing school value-added in the context of Mexico. However, there are two disadvantages of ENLACE data: first, the questionnaire is given only to a random subset of tested students. Second, we cannot access ENLACE performance data for students enrolled in UNAM, since UNAM did not participate in the ENLACE assessment.

The last piece of data comes from administrative records on teacher salaries. In Mexico, there is no centralized source that reports teacher salaries at the school level, as compensation is managed separately by different educational subsystems. To overcome this limitation, we submitted information requests through Mexico's Transparency Portal to the main subsystems operating in Mexico City and the State of Mexico. The responses provide payroll records at the individual level, which, once harmonized across subsystems, allow us to link teachers to specific schools. This linkage makes it possible to construct school-level measures of teacher salaries. For the purpose of this study, we focus on the 2018 school year, calculating the average gross monthly teacher salary at the school level. While contractual arrangements and pay scales differ across subsystems, these data provide a unique opportunity to compare salaries in the COMIPEMS region.[8]

The main year for our analysis is 2019 for two reasons: first, it is the last year before the onset of the pandemic; second, the introduction of Beca Universal de Educación Media Superior Benito Juárez Scholarship (BJS) unifies and replaces a variety of pre-existing universal cash transfer programs, such as Prepa Sí, that previously targeted high school students in Mexico City. Before BJS, students benefited from multiple overlapping schemes that are not consistently identified in our data, making it difficult to separate the program effects of admission outcomes. By focusing

[7] A z-score measures how many standard deviations a value lies from the mean of its distribution and is calculated as $z = \frac{x-\mu}{\sigma}$, where x is the individual score, μ is the population mean, and σ is the population standard deviation.

[8] Importantly, federal subsystems such as DGETI and DGETAyCM, as well as some state subsystems, including *Preparatorias Oficiales* and *Telebachilleratos*, did not respond positively to our requests. This was mainly due to the failure of transparency units, particularly those under the SEP, to comply and provide the requested information. Additional steps are currently being taken to obtain these data through Mexico's oversight authority, since teacher salaries are considered public information.

on 2019, we avoid these confounding factors and capture the last cohort before both the pandemic shock and the consolidation of transfer programs. Lastly, because the ENLACE data are from 2013–2018, we assume that school quality remains stable over time for estimation purposes.[9]

The 2019 COMIPEMS dataset includes 310,159 applicants. For further sample selection, we begin by excluding applicants with inconsistent middle school information or invalid postal codes, which account for less than 5% of the full sample. We retain only applicants aged 15–16 and restrict to current-year graduates, defined as those who obtained their middle school certificate in 2019. We then exclude applicants without a certificate, those who did not take the exam, and those disqualified for infractions. Finally, we remove applicants residing outside Mexico City and the State of Mexico.

Table 2 summarizes 172,625 COMIPEMS applicants in 2019 after sample selection. The cohort is 51% female. Parental schooling is modest: Mothers have an average of 6.7 years of education, and household heads have an average of 7.2 years. Households are large, with about five members on average; 54% of applicants reside in Mexico City, and 37% report English fluency. Economic conditions vary widely: monthly income per capita averages 1,481 MXN with a standard deviation of 1,286, and about 66% of applicants come from households below the extreme food poverty line, indicating general economic disadvantage in the sample. Academic performance is strong but varies considerably across students. The average 9th grade GPA is 8.23 out of 10, indicating that a large share of students meet the GPA threshold for elite schools. The placement test has a mean score of 68.27 out of 128 with a standard deviation of 21.48, and the average GPA of admitted students in middle school typically ranges between 8.1 and 8.3.

Students submit long, diversified preference lists, averaging about ten options across nearly ten distinct schools. On average, each list includes 2.66 elite academic programs from UNAM, 2.83 non-elite academic programs, 1.57 elite technological programs from IPN, 1.87 non-elite technological programs, and 0.87 vocational programs. On average, students apply to their first-choice school located about 12.6 km from home, with substantial variation across applicants (SD 11.0 km). The farthest school on their submitted preference lists is, on average, 24.5 km away, indicating that while many students prioritize nearby options, a considerable share are willing to consider schools located much farther from their residence. Overall admission rate is 92%. Placements are 36% in non-elite academic, 24% in non-elite technological, 10% in elite academic, 8%

[9] ENLACE was discontinued after 2018 and replaced by PLANEA. We plan to obtain Nacional para la Evaluación de los Aprendizajes (PLANEA) data in future work to obtain an updated measure of school quality.

in elite technological, and 14% in vocational programs. Taken together, the data show substantial socioeconomic and academic heterogeneity, broad aspirations for elite programs, and a placement pattern consistent with tight selectivity at elite schools. The average distance between a student's home postal code and the school to which he/she is admitted is 10.1 km, indicating that most students enroll in relatively nearby schools.

Table 3 reports school-level means with standard deviations by school type, defined as the combination of track and elite status, and shows large and systematic differences across types. Unless otherwise noted, "type" refers to the combination of track and elite status for the remainder of the paper. Elite academic (UNAM) and elite technological (IPN) schools enroll the highest-achieving cohorts, with middle-school GPAs around 8.9 and 8.7 and placement scores near 103 and 95. Non-elite academic, non-elite technological, and vocational schools serve weaker cohorts, with placement scores typically in the mid-50s to low-60s for the former and low-50s for vocational. Academic tracks have higher female shares, roughly between 0.51 and 0.54, while technological tracks range from about 0.39 to 0.48. Student–teacher ratios are smaller in UNAM and IPN, averaging about 12 to 13 students per teacher, compared with their non-elite counterparts. This suggests that elite schools are able to maintain smaller class sizes, which is typically associated with higher instructional quality and greater individual attention for students. Average net salaries are also graded, with UNAM the highest and vocational the lowest. Regarding spatial access, elite schools are located in dense, transit-rich areas with more public transit stops within one to three kilometers, and the nearest public transit is within 1km away. In contrast, non-elite schools are more dispersed and at a greater distance from the nearest transit. Hardware stocks are larger in UNAM and IPN, with hundreds of computers and extensive classroom use, compared to non-elite schools. Lastly, vocational and non-elite technological schools offer more specialties per campus, meaning they provide a wider range of programs or majors such as accounting, electronics, mechanics, or computer science within the same institution, whereas UNAM, IPN, and non-elite academic schools are largely single-track, focusing on a uniform academic curriculum.

4 Value-added and Student Sorting Pattern

In this section, we describe how we estimate school value-added, which serves as our measure of public school quality. We then document the sorting of students under the SD mechanism, focusing on the correlation between school value-added with admission regarding student achieve-

Table 2: Summary Statistics of COMIPEMS Applicants (Year 2019)

	Mean	Std. Dev.	Min	Max
Sociodemographics				
Female	0.51	0.50	0	1
Mother's Years of Education	6.74	3.04	1	14
Household Head's Years of Education	7.15	3.32	1	15
Household Size	4.99	1.65	1	9
Residence in Mexico City	0.54	0.50	0	1
Fluent in English	0.37	0.48	0	1
Monthly Household Income per Capita ($MXN)	1,480.68	1,285.90	0	22,531.50
Below Extreme Food Poverty Line	0.66	0.47	0	1
Middle School Information				
9th Grade GPA (out of 10)	8.23	0.87	6	10
Graduated from General Academic Middle School	0.64	0.48	0	1
Graduated from Technological Middle School	0.30	0.46	0	1
Middle School Average Math GPA of Admitted Students	8.10	1.12	6	10
Middle School Average Spanish GPA of Admitted Students	8.30	1.12	6	10
Performance in Placement Test				
Total Score (out of 128)	68.27	21.48	8	128
ROL Submission				
# Options	10.02	4.23	1	20
# Schools	9.80	4.18	1	20
# Elite Academic (UNAM) Schools	2.66	2.63	0	14
# Non-elite Academic Schools	2.83	2.29	0	20
# Elite Technological (IPN) Schools	1.57	2.08	0	16
# Non-elite Technological Schools	1.87	1.92	0	20
# Vocational Schools	0.87	1.23	0	18
Distance (km) to the First Applied School	12.63	10.95	0	199.78
Max. Distance (km) to Submitted Schools	24.45	14.36	0	235.06
Admission Outcome				
Admitted	0.92	0.27	0	1
Admitted to Elite Academic (UNAM) Schools	0.10	0.30	0	1
Admitted to Non-elite Academic Schools	0.36	0.48	0	1
Admitted to Elite Technological (IPN) Schools	0.08	0.27	0	1
Admitted to Non-elite Technological Schools	0.24	0.43	0	1
Admitted Vocational Schools	0.14	0.34	0	1
Distance (km) to Admitted School	10.13	10.42	0	228.94
# Students				172,625

Notes: The table reports summary statistics for applicants to the COMIPEMS admissions process in 2019. Variables are grouped into: (i) sociodemographics, (ii) middle school information, (iii) performance on the placement test, and (iv) variables related to the rank-order list (ROL), for example, the number of options and schools submitted, and (v) admission outcomes. The sample includes 172,625 applicants. For each variable, the table reports the mean, standard deviation, minimum, and maximum. Monthly household income per capita is expressed in Mexican pesos (MXN).

Table 3: Summary Statistics of High Schools in COMIPEMS

	Elite Academic (UNAM)		Non-elite Academic		Elite Technological (IPN)		Non-elite Technological		Vocational	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Panel A: Information on Admitted Students (2018)										
% Female Students	0.51	(0.03)	0.54	(0.08)	0.39	(0.14)	0.48	(0.09)	0.46	(0.10)
% Indigenous Students	0.02	(0.01)	0.06	(0.06)	0.02	(0.01)	0.05	(0.02)	0.06	(0.02)
% Students from Private Middle School	0.25	(0.10)	0.03	(0.04)	0.17	(0.07)	0.03	(0.02)	0.02	(0.01)
% Students from General Academic Middle School	0.69	(0.05)	0.64	(0.16)	0.64	(0.04)	0.63	(0.07)	0.62	(0.07)
Average Middle School GPA	8.90	(0.16)	7.98	(0.31)	8.67	(0.15)	8.04	(0.20)	7.78	(0.15)
Average Placement Test Score	102.96	(6.27)	56.43	(12.71)	94.72	(5.69)	63.19	(8.56)	52.92	(6.23)
Panel B: Information on Teachers (2018)										
% Female Teachers	0.51	(0.05)	0.53	(0.17)	0.44	(0.07)	0.47	(0.11)	0.47	(0.12)
Student-teacher Ratio	11.73	(3.99)	29.26	(64.66)	13.30	(7.90)	14.46	(9.95)	11.02	(5.63)
Monthly Average Gross Salary	22197.35	(1912.93)	14216.31	(1995.76)	18343.70	(2422.81)	13063.99	(2947.54)	7576.35	(1395.24)
Monthly Average Net Salary	16555.98	(1356.08)	10321.71	(1304.17)	11704.79	(1748.72)	8994.88	(2519.85)	5877.36	(717.13)
Panel C: Geographical Information										
# Nearby Schools within 5km	12.64	(6.17)	18.52	(12.07)	14.81	(6.10)	15.05	(9.74)	15.14	(8.89)
# Nearby Schools within 7km	26.36	(10.69)	32.62	(18.45)	30.00	(10.45)	27.55	(15.52)	28.65	(15.43)
Distance to the Nearest School	1.23	(0.85)	1.14	(0.99)	1.24	(0.78)	1.15	(0.87)	1.21	(1.58)
Average Distance to the Nearest 3 Schools (km)	1.82	(0.69)	1.73	(1.20)	1.86	(0.83)	1.80	(0.99)	1.80	(1.71)
Average Distance to the Nearest 5 Schools (km)	2.32	(0.67)	2.19	(1.45)	2.36	(0.95)	2.29	(1.15)	2.30	(1.86)
# Public Transit Stops within 1km	16.57	(13.86)	1.95	(4.69)	10.75	(14.74)	4.23	(9.89)	2.93	(4.48)
# Public Transit Stops within 2km	55.57	(41.47)	8.18	(13.18)	49.75	(49.30)	17.60	(35.25)	14.89	(16.59)
# Public Transit Stops within 3km	123.14	(85.55)	18.86	(26.97)	111.06	(84.69)	37.55	(63.87)	32.72	(36.93)
Distance to the Nearest Public Transit (km)	0.42	(0.29)	1.96	(2.35)	0.64	(1.01)	1.54	(1.47)	1.06	(0.77)
Average Distance to the Nearest 3 Public Transit (km)	0.55	(0.28)	2.49	(2.83)	0.78	(1.22)	1.95	(1.85)	1.41	(1.29)
Average Distance to the Nearest 5 Public Transit (km)	0.63	(0.30)	2.81	(3.14)	0.93	(1.41)	2.21	(2.08)	1.62	(1.52)
Panel D: Resource on Computers										
# Computers	733.50	(167.23)	52.35	(68.81)	761.38	(187.47)	149.32	(104.62)	180.14	(47.69)
# Computers in Operation	733.29	(167.31)	47.18	(64.86)	745.00	(180.89)	133.80	(94.54)	167.65	(45.98)
# Computers for Education Purpose	607.93	(209.48)	33.10	(44.37)	526.12	(161.96)	104.42	(76.78)	127.35	(43.09)
# Computers for Teaching Purpose	54.36	(47.93)	2.65	(6.90)	102.75	(66.20)	4.62	(6.01)	4.56	(2.37)
# Computers for Administrative Purpose	71.00	(76.05)	11.43	(18.35)	116.12	(34.94)	24.76	(18.36)	35.74	(10.77)
# Computers for Education Use in Classroom	471.21	(269.16)	31.48	(41.66)	465.50	(226.37)	99.74	(76.14)	119.63	(44.92)
Panel E: Specialty										
# Specialties	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.88	(1.18)	3.32	(0.63)
# Schools	14		266		16		121		57	

Notes: This table reports school-level summary statistics for high schools participating in COMIPEMS. Schools are grouped into five categories: elite academic (UNAM), non-elite academic, elite technological (IPN), non-elite technological, and vocational. Variables are organized into five panels: (A) admitted student composition in 2018 (e.g., % female students, share from private middle schools, average middle school GPA, average placement test score), (B) teacher characteristics at the school (e.g., % female teachers, student–teacher ratio, monthly average gross and net salary), (C) school geography (e.g., distance to the nearest high school; average distance to the nearest 3 and 5 schools; number of public transit stops within 1–3 km; distance to the nearest public transit), (D) computer resources (e.g., total computers, computers in operation, computers for teaching or administration, computers used in the classroom for education use), and (E) number of specialties offered. For each variable, the table reports the mean and standard deviation across schools within each category. Teacher salaries are expressed in Mexican pesos (MXN), and distances are measured in kilometers (km). The final row reports the number of schools in each category.

ment and demographics.

4.1 Value-added of High Schools in COMIPEMS

We estimate school quality by applying a time-invariant value-added framework with Empirical Bayes (EB) shrinkage, following the approach widely used in the value-added literature (e.g., Meyer, 1997; Ladd and Walsh, 2002; Kane and Staiger, 2008; Chetty et al., 2014; Jackson et al., 2020; Angrist et al., 2023, 2024). More specifically, academic achievement of student i at school s in year t, y_{ist}, is expressed as:

$$y_{ist} = X'_{ist}\beta + \theta_s + \delta_t + \varepsilon_{ist}, \tag{1}$$

where X_{ist} denotes the vector of controls for student demographics, parental characteristics, and academic ability, measured by prior academic performance.[10] δ_t is a fixed effect for year t. θ_s serves as the raw time-invariant value-added in our model, and ε_{ist} is an idiosyncratic test score shock, which is assumed to be i.i.d. $N(0, \sigma^2_\epsilon)$. The value-added measures school quality by isolating the contribution of each school to student academic outcomes, after accounting for student background and ability.

The analysis relies on student achievement data from the ENLACE examinations, a nationwide standardized test administered annually in Mexico from 2006 to 2013 to all students in their final year of upper secondary education. Alongside the test, schools submit administrative forms containing detailed background information, including student demographics, parental education, and other contextual variables. Because ENLACE is administered only once at the end of upper secondary school, it does not permit the construction of traditional value-added measures that rely on gains between baseline and follow-up test scores. Instead of conditioning on lagged achievement, we estimate school contributions by conditioning end-of-high-school performance on a rich set of student and family controls (e.g., Vivanco, 2013; Cunha and Miller, 2014; Navarro and Pariguana, 2025). This specification allows us to isolate school contributions without relying on prior test scores, drawing on ideas from the contextual value-added literature.

To improve reliability, we apply an Empirical Bayes shrinkage procedure, which adjusts the

[10] Specifically, X_{ist} in Equation (1) includes: (i) student demographics such as gender, age, working status, marital status, and interaction between gender and marital status; (ii) elementary school performance (school type, GPA, and interaction with gender); (iii) middle school performance (school type, modality, GPA, and gender interaction); (iv) high school indicators including merit- and need-based scholarship receipt, current GPA, and interaction with gender; and (v) parental education.

fixed effect estimates toward the overall mean based on their precision (Kane and Staiger, 2008; Chetty et al., 2014; Koedel and Rockoff, 2015; Herrmann et al., 2016). This correction is especially important for small schools with limited student observations, where raw estimates are highly variable. The resulting shrunken value-added estimates better capture the true contribution of the school to student achievement net of family background and other controls. Further details on the implementation of the EB shrinkage procedure for estimating value-added are presented in Appendix B.1.

Before we present the estimates of value-added, we first want to discuss two important considerations regarding our measurement of school quality. First, we use value-added in mathematics as our primary measure of school quality rather than value-added in Spanish. Mathematics achievement has been shown to be more predictive of long-term outcomes such as college enrollment, labor market earnings, and performance on international assessments, thereby providing stronger external validity for assessing school effectiveness (Adelman, 1999; Rose and Betts, 2004; Long et al., 2009; Joensen and Nielsen, 2009; Goodman, 2019). Moreover, mathematics tests are typically more standardized and less sensitive to curricular or cultural variations across schools, which improves comparability. For these reasons, we consider mathematics value-added a cleaner and more policy-relevant proxy for school quality, while acknowledging that language value-added may capture complementary dimensions of student development. Second, as mentioned in Section 3, UNAM does not participate in the ENLACE assessment. To approximate the quality of UNAM high schools, we collect detailed school characteristics and construct predictive measures of performance based on observable attributes.

Figure 1 plots kernel densities of school-level value-added in math ENLACE z-score by school type, with the scale normalized so that zero equals the sample average school effect conditional on student covariates and year fixed effects; positive values indicate above average contributions to achievement and negative values indicate below average contributions.

Panel 1a shows the elite distribution tightly concentrated well to the right, while the non-elite distribution centers near zero with a much larger spread. This location shift indicates systematically higher math productivity in elite schools rather than differences in observed student composition. The limited overlap indicates that, for students, moving from a typical non-elite school to an elite school would raise expected math achievement. At the same time, the remaining overlap shows that a nontrivial subset of non-elite schools performs on par with the lower end of the elite distribution. Panel 1b shows vocational schools are concentrated below zero with little right tail, consistent with weak math training. General academic and technological schools cluster around zero with modest right tails, implying that the median non-elite school neither adds nor subtracts much in math relative to the average school, but there is meaningful upside in the upper quantiles. Appendix B.2 provides the estimation table of value-added for both Spanish and math. Appendix B.3 reports the distribution of estimated school value-added on ENLACE Spanish.

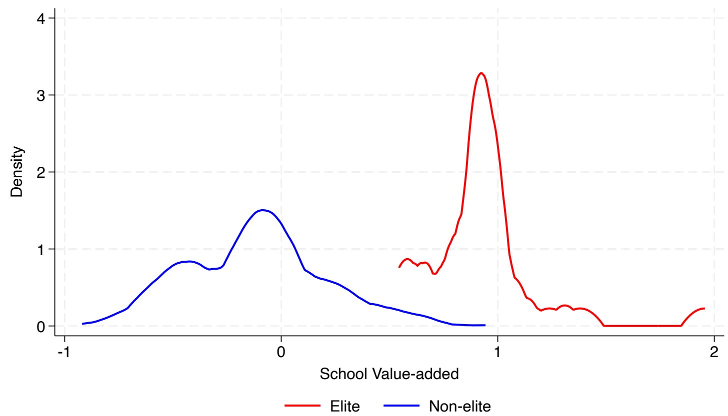
To better interpret the magnitudes, we define scaled (rescaled) value-added as the estimated value-added (in z-score units) multiplied by the national standard deviation of ENLACE test scores, 100. This places value-added on the same scale as ENLACE scores: for instance, a scaled value-added of 100 means that, holding other factors constant, students attending that school score on average 100 points higher on the ENLACE exam than they would have without the school's contribution. Throughout the paper, we refer to this measure as the scaled (or rescaled) value-added.

4.2 Data Pattern: Student Sorting

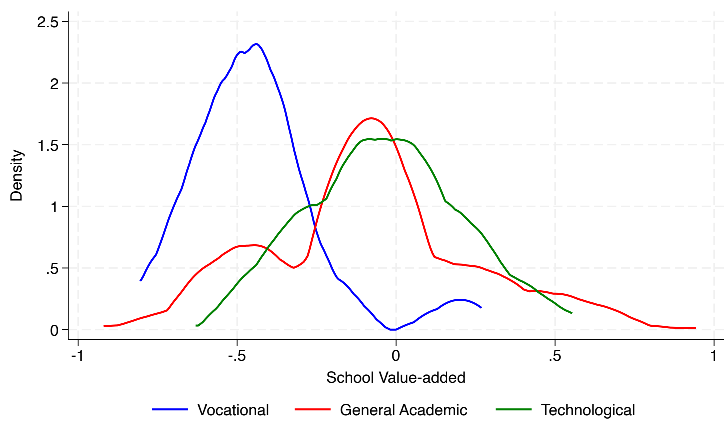
In this section, we document sorting under the admission rules before 2025. We relate school quality, measured by math value-added, to each school's admission cutoff, defined as the test score of the last admitted student and used as an indicator of selectivity. We also examine how value-added correlates with the test scores and demographic composition of admitted students. More specifically, we show correlations between value-added and (i) the admission cutoff and the within-school dispersion of admitted students' placement test scores, and (ii) the composition of admitted students regarding income, poverty, and gender. We use these patterns to show that assignment rules shape admission outcomes and determine who is placed where. In turn, they affect access to high value-added schools and the distribution of achievement, socioeconomic composition, and segregation across schools.

Figure 2 maps school math value-added to two measures of admissions selectivity, plotting school-level scatters with separate fitted lines for elite and non-elite schools. Panel 2a plots school math value-added on the x-axis against the admission cutoff in the placement test on the y-axis. For both elite schools and non-elite schools, the slopes are positive. Higher value-added schools tend to require higher placement scores for admission, consistent with the outcome under the SD mechanism with test scores as priority rules. Elite schools are concentrated at both high value-added and high cutoffs. Panel 2b relates math value-added to the within-school standard deviation of admitted students' placement scores. For non-elite schools, the fitted line slopes downward: higher value-added schools draw more homogeneous cohorts regarding academic performance, consistent with tighter selection around the cutoff. Elite schools cluster within a narrow range of admission scores and exhibit consistently high value-added.

Figure 1: Distribution of Estimated School Value-added on ENLACE Math by School Type



(a) Elite vs. Non-elite

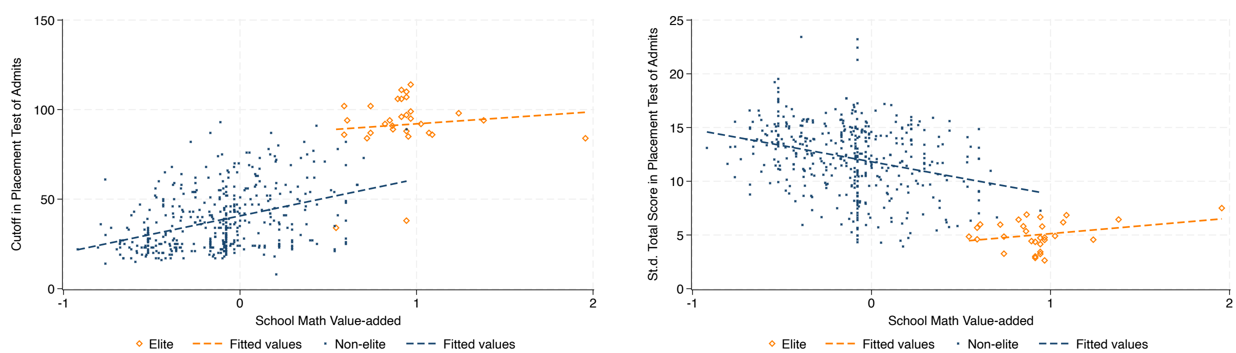


(b) Non-Elite Schools by Track

Notes: This figure shows the kernel density of estimated school value-added in ENLACE math by school type. Panel (a) compares elite and non-elite schools; Panel (b) decomposes non-elite schools by track (technological, general academic, and vocational).

Figure 3 plots math value-added against the socioeconomic composition of admitted students. Panel 3a shows a positive association between value-added and the share of admits from high-income households in both elite and non-elite sectors.[11] Elite schools admit a larger share of high-income students than non-elite schools; the cross-school variation is limited within elite schools. Panel 3b shows that a higher value-added school is associated with a lower share of admits from households below the extreme food poverty line. These correlations indicate that schools with higher value-added tend to enroll less disadvantaged cohorts. Figure 4 relates school math value-added to the share of female admits, separately for elite and non-elite schools. Among non-elite schools, the fitted line is essentially flat: higher value-added is not systematically associated with a higher or lower fraction of female students, and the average female share is around one-half. For elite schools, the fitted line slopes modestly downward; schools with higher value-added tend to have slightly lower female shares. This pattern is consistent with the data: roughly one-half of applicants are female, and the assignment mechanism does not condition on gender.

Figure 2: School Value-added and Selectivity of Admitted Students (by Elite Status)



(a) Admission Cutoff vs. Math Value-added

(b) Within-school Dispersion of Admitted Students' Placement Scores vs. Math Value-added

Notes: This figure plots school math value-added against two measures of admission selectivity, with separate scatters and fitted lines for elite and non-elite schools. Panel (a) shows the correlation between admission cutoffs in the placement test and value-added. Panel (b) relates value-added to the within-school dispersion of admitted students' placement test scores. Each dot represents one school.

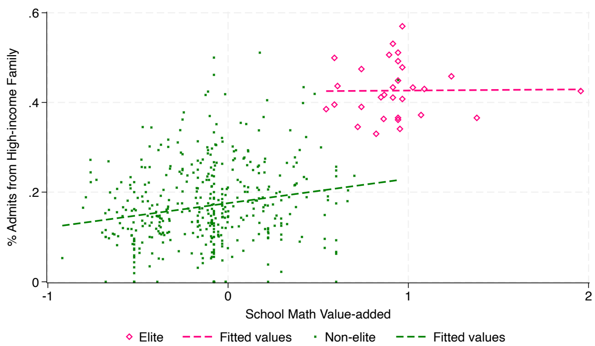
5 Empirical Model

In this section, we introduce our empirical model of student demand and school quality supply. Our framework links student preferences to demand and the assignment rule, and models schools' quality choices from enrollment incentives and costs. We use it to recover preference parameters and supply-side parameters in the school quality function and to assess how changes in

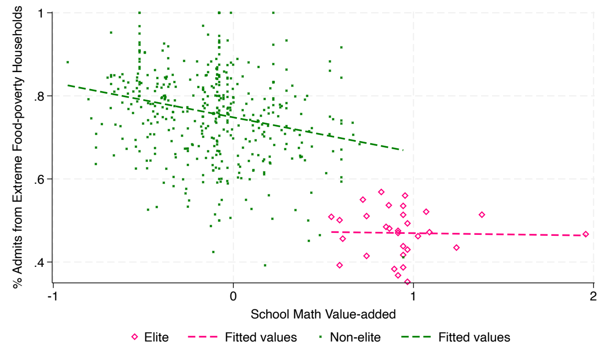
[11]High-income households are defined as the top 25% of the sample income distribution.

assignment rules affect enrollment and school quality.

Figure 3: School Value-added and Socioeconomic Composition of Admitted Students (by Elite Status)



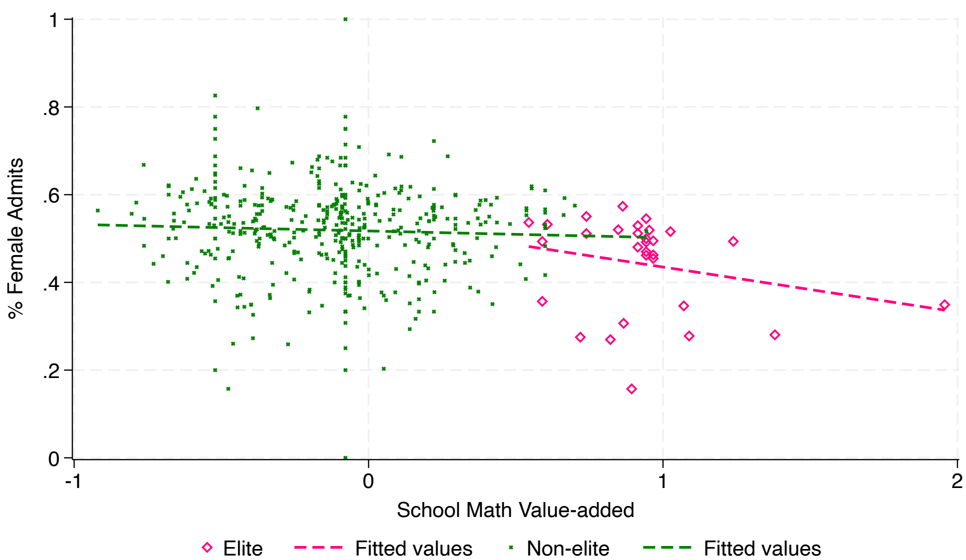
(a) Percent Admitted from High-income Households vs. Math Value-added



(b) Percent Admitted from Below Extreme Food Poverty Line Households vs. Math Value-added

Notes: This figure relates school math value-added to the socioeconomic composition of admitted students, with scatters and fitted lines for elite and non-elite schools. Panel (a) shows the relationship between school value-added and the share of students from high-income households. Panel (b) shows the relationship between school value-added and the share of students from below extreme food poverty line households. Each dot represents one school.

Figure 4: Math Value-added and Share of Female Admits (by Elite Status)



Notes: This figure relates school math value-added to the share of female admits, with scatters and fitted lines for elite and non-elite schools. Each dot represents one school.

5.1 Student Demand

On the demand side of the model, student i chooses school j in order to maximize his/her indirect utility:

$$u_{ij} = \beta^q q_j + \beta^x x_i + \gamma \log d_{ij} + \beta^Q Q_j + \xi_j + \epsilon_{ij} \tag{2}$$

where q_j is school j's quality, i.e., estimated value-added on math ENLACE, d_{ij} is the driving distance (in kilometers) from student i's residence postal code to school j, x_i is a vector of student characteristics, which include an indicator for female, whether the student is a resident of Mexico City, the monthly household income per capita, 9th grade GPA, and the mother's years of education. The variable Q_j is a vector of observed school attributes, including the percentage of female teachers, the student-teacher ratio, the percentage of admitted female students, the percentage of admitted students who graduated from private middle schools, the percentage of admitted students who graduated from general middle schools, the average middle school GPA of admitted students, whether the school j is in Mexico City, the number of specialties provided together with a track fixed effect. The term ξ_j is a structural error term, which is interpreted as the unobserved school quality or amenity that is valued in common by students, whereas ϵ_{ij} is an error term that follows a type-I extreme distribution. Lastly, we assume student i's utility from the outside option to be $u_{i0} = \epsilon_{i0}$, which follows a type-I extreme distribution as well. We normalize the mean utility of not being assigned to any school to be zero. Note that the outside option combines any type of educational choice other than COMIPEMS schools, such as private schools or homeschooling.

The parameters to be estimated are β^x, γ, β^q. The coefficient vector β^x measures how the observed student characteristics shift the systematic preference of choosing schools in COMPIEMS relative to the outside option. γ measures sensitivity to distance, which proxies for commuting costs. On the school side, β^q reflects the importance students place on value-added when choosing schools, and β^Q captures the role of other observed quality attributes. Together, these parameters quantify how both student characteristics and school attributes shape demand.

It is helpful to re-write Equation (2) as:

$$u_{ij} = \delta_j + \beta^x x_i + \gamma \log d_{ij} + \epsilon_{ij}, \tag{3}$$

where δ_j is the mean utility, absorbing school-level observed and unobserved components that influence students' utility,

$$\delta_j = \beta^q q_j + \beta^Q Q_j + \xi_j. \tag{4}$$

This expression is particularly useful for estimation. Equation (3) provides the basis for deriving individual choice probabilities and aggregate demand, while Equation (4) makes explicit that the unobserved component ξ_j may be correlated with observed school quality q_j, thereby raising an endogeneity concern. We detail in the later section why endogeneity arises and how we address it to consistently estimate β^q. This parameter is central to our analysis, as it governs the elasticity of demand with respect to value-added, which in turn plays a critical role in the supply-side model.

5.2 School Supply

On the supply side, we follow Gilraine et al. (2023) and model public high schools to set their educational quality as value-added. Specifically, public school j chooses its value-added q_j based on the following function:

$$F_j\left(D_j(\mathbf{q})\right) + G_j\left(q_j\right) - C_j(\mathbf{q}), \tag{5}$$

where $C_j(\mathbf{q}) = mc_j\left(q_j\right) D_j(\mathbf{q}) + \eta q_j$. The "revenue" of the school setting q_j comes from two parts. First, $F_j(\cdot)$ is a function representing how public schools value total student enrollment, which is represented by $D_j(\mathbf{q})$. Note that demand for school j depends not only on its own value-added but also on the value-added of competing schools, since students choose among all available options and relative quality determines enrollment shares. Second, $G_j(\cdot)$ represents how a school benefits from higher quality that is not driven by demand, such as higher prestige or intrinsic value. The cost of providing quality, $C_j(\mathbf{q})$, is also driven by two parts. The first part is the variable cost, which is the multiplication of marginal cost $mc_j\left(q_j\right)$ and demand. Note that it is the marginal cost per student enrollment and depends on school j's own value-added. The cost of providing education rises with both the number of enrolled students and the level of quality the school offers.The second part ηq_j captures the fixed cost of maintaining value-added that is independent of enrollment. It is the cost the school incurs to raise its value-added, regardless of the number of students it has, including expenses such as curriculum, infrastructure, or technology.

We assume a linear marginal cost function, $mc_j\left(q_j\right) = \pi_j + \kappa q_j$, where π_j is the fixed marginal cost of educating one student at zero quality (per pupil expenditure), and κ is common across schools and reflects how costly it is to raise quality per student. If $\kappa > 0$, the cost increases with q_j, which reflects diminishing returns or increasing difficulty in getting higher quality. In the empirical model, we further parameterize π_j as a function of observed local cost, proxied by teacher wages, and include a measurement error term. Thus, the marginal cost function becomes $mc_j(q_j) = \beta X_j + \kappa q_j + \omega_j$.

To characterize schools' optimal quality decision, we derive the first-order condition with respect to q_j in Equation (5), yielding the optimal level of value-added given the demand implied by students' choices and the quality of competing schools. Following Gilraine et al. (2023) and Neilson (2021), we rewrite the first-order condition in a more intuitive form:

$$q_j^* = \underbrace{\left(\frac{F_j'\left(D_j\left(\mathbf{q}^*\right)\right)}{\kappa} - \frac{\beta T_j}{\kappa} - \frac{\omega_j}{\kappa}\right)}_{\text{Value-added under Perfect Competition}} - \underbrace{\left(1 - \frac{G_j'\left(q_j^*\right) - \eta}{\kappa} D_j\left(\mathbf{q}^*\right)^{-1}\right)\frac{1}{\sigma_j\left(\mathbf{q}^*\right)}}_{\text{Value-added ``Markdown''}}, \tag{6}$$

where $\sigma_j(\mathbf{q}) = \frac{1}{D_j(\mathbf{q})}\frac{\partial D_j(\mathbf{q})}{\partial q_j}$ denotes school j 's own-value-added semi-elasticity of demand.[12] It measures the percentage change in school j's demand associated with a one-unit change in its own value-added, holding other schools' quality levels fixed; that is, the sensitivity of student enrollment to marginal improvements in quality relative to its current demand level.

Equation (6) expresses the school's optimal value-added choice as the difference between two components. The first bracket is interpreted as the *Value-added under Perfect Competition*. It corresponds to the quality level that would prevail if schools acted as price-takers, treating enrollment as given and ignoring the effect of their own quality on demand. The second bracket is the *Value-added "Markdown"*, which captures how schools internalize their own demand curve when choosing quality. In contrast to the price-taking benchmark, here schools recognize that raising value-added affects both costs and enrollment.

The size of this markdown is governed by the own-value-added semi-elasticity $\sigma_j(\mathbf{q})$. When demand is more responsive, competitive pressure is stronger and the markdown is smaller, leading schools to supply higher quality. Conversely, when a school faces limited competition—effectively operating as a local monopolist—the markdown is larger, reducing the incentive to improve quality. In our setup, the markdown term also incorporates direct incentives or constraints on quality provision through $G_j'(q_j)$ and η, which shift the coefficient on the demand semi-elasticity relative to the standard profit maximization case.

[12] See Appendix C for a detailed derivation.

6 Identification and Estimation

The empirical model is estimated in several steps. First, we estimate student preferences based on a two-sided matching framework to recover demand parameters and mean utilities. Second, we discuss the identification and estimation of the endogenous parameter, β^q, i.e., the student preferences on value-added. Lastly, we further discuss the supply model specification and estimation, including demand shifters that address the endogenous concerns of demand.

6.1 Demand Estimation

As discussed in Section 2.1, before 2025, COMIPEMS assigned students using a centralized mechanism based on test scores and students' ROLs. This SD mechanism is a special case of the Gale–Shapley deferred acceptance (DA) mechanism that arises when all schools share a common strict priority ordering over students. While DA is considered strategy-proof (Abdulkadiroğlu and Sönmez, 2003), it is well documented in the literature that assuming students submit ROLs that reveal their true preferences over schools under DA is restrictive (e.g., Haeringer and Klijn, 2009; Fack et al., 2019; Artemov et al., 2023). Therefore, leveraging ROLs as true preference ordering for estimation may generate implausible estimates. This is especially true if students' ROLs are restricted or if students face only limited uncertainty about their admission outcomes. For example, in the context of COMIPEMS, a student who prefers elite schools but has a low test score may choose not to apply to there, which implies that not all students have incentives to rank all schools truthfully in their ROLs.

We estimate preferences using the Stability assumption proposed by Fack et al. (2019). The assumption states that the observed matching between students and schools is stable in equilibrium. This stability-based approach has been used in several matching settings (e.g., Akyol and Krishna, 2017; Combe et al., 2022; Ainsworth et al., 2023; Bobba et al., 2021, 2023), and has been shown to produce more precise estimates than approaches that assume (weak) truth-telling in students' submitted ROLs of limited length. Fack et al. (2019) also show that the Stability assumption leads to a discrete choice model with personalized choice sets. Consequently, identification follows standard arguments from the discrete choice literature (e.g., Matzkin, 1993).

In our setting, under the Stability assumption, every applicant is matched with his/her most preferred school among those that are feasible. Mathematically, denote student i's score as s_i, the ex post cutoff of school j as c_j.[13] School j is feasible to student i if $s_i > c_j$. The set of feasible schools is called the personalized choice set, denoted as $\mathcal{C}\left(s_i,c_j\right)$. Suppose school j is assigned to student i in her personalized choice set. Further, denote $\mathcal{A}\left(s_i, c_j\right)$ as the union of outside option and personalized choice set. We denote the outside option as O, then $\mathcal{A}\left(s_i, c_j\right) = \{O\} \cup \mathcal{C}\left(s_i, c_j\right)$. Her preference becomes:

$$u_{ij} > u_{ij'}, \text{ where } j' \neq j \text{ and } j' \in \mathcal{A}\left(s_i, c_j\right). \tag{7}$$

Equation (7) indicates that the enrolled students prefer the assigned school over all other options, including those in their personalized choice set and the outside option.

Regarding estimation, we use maximum likelihood in a multinomial logit model. The probability that student i chooses school j in their personalized choice set is given by:

$$\begin{aligned} \mathrm{Pr}_{ij} &= \mathrm{Pr}\left(u_{ij} > u_{ij'}, \text{ where } j' \neq j \text{ and } j' \in \mathcal{A}\left(s_i, c_j\right)\right) \\ &= \mathrm{Pr}\left(\underset{j \in \mathcal{A}(s_i, c_j)}{\arg\max}\ u_{ij} \mid \delta_j, x_i, d_{ij}, \mathcal{A}\left(s_i, c_j\right)\right) \\ &= \frac{exp(\delta_j + \beta^x x_i + \gamma \log d_{ij})}{\sum_{j \in C(s_i, c_j)} exp(\delta_j + \beta^x x_i + \gamma \log d_{ij}) + 1}. \end{aligned} \tag{8}$$

The log-likelihood function is then defined as:

$$L(\beta) = \sum_i log(Pr_{ij}), \tag{9}$$

where β represents the vector of demand parameters, which includes the vector of mean utilities δ, the vector of coefficients on the student characteristics β^x, as well as student tastes for distance γ. We then back out the choice probability and demand with the estimated demand parameters.

Estimating a large-scale discrete choice model under personalized choice sets with hundreds of school options and tens of thousands of students requires both efficient computation and flexible model specification. Conventional econometric packages become infeasible in this setting because

[13] Cutoffs are well-defined and often observable to the researcher: given the admission outcome, each school's cutoff is the lowest priority index of the students accepted here. In our case, the lowest score corresponds to the score of the last admitted student.

of the high-dimensional choice set and the need to evaluate millions of choice probabilities. We use the new Python package, Torch-Choice, developed by Du et al. (2023), to perform large-scale choice modeling. The suitability of Torch-Choice for our estimation is based on several salient features. First, built on PyTorch, it facilitates GPU acceleration and employs a custom data structure, thereby enhancing computational efficiency and conserving memory substantially. Second, it is capable of accommodating large datasets, enabling our computation with a vast number of applicants and options. Third, its design allows users considerable flexibility in specifying the desired functional form as well as feasible options.

6.2 Identification of β^q

As mentioned in Section 5.1, a central concern in estimating β^q is that school value-added q_j may be correlated with the unobserved component ξ_j. Conceptually, schools that appear to generate high value-added might simultaneously differ in unobserved attributes that are not fully captured by observed controls, such as reputation, safety, and leadership. To address endogeneity, we use an instrument based on the 2018 weighted average of classroom computer availability at competing schools. This instrument shifts the competitive pressure faced by a school while remaining plausibly unrelated to its unobserved demand shocks. Specifically, the instrument is defined as

$$z_j = \sum_{k \neq j} w_{jk} \cdot \text{ \# computers}_k,$$

where the summation is taken over all other schools $k \neq j$ within the same type (track, elite status), # computers_k refers to the number of computers in the classroom used at school k. The weights w_{jk} are constructed from relative market shares,

$$w_{jk} = \frac{\text{\# students}_k}{\sum_{l \neq j} \text{\# students}_l},$$

so that schools with larger enrollments exert greater influence on the instrument.

Before discussing the validity of the instrument, note that our IV of computer resources is constructed using 2018 data, which predates the 2019 COMIPEMS application and school year. This timing makes the instrument predetermined: It is fixed prior to the 2019 admissions cycle and teaching year and therefore cannot be affected by 2019 demand shocks. Now, we turn to discuss instrument validity:

(i) **Relevance:** Rival schools' computer resources measured in 2018 are expected to shift a school's 2019 value-added through competitive pressure: when nearby rivals are better equipped, schools adjust practices to raise value-added.

(ii) **Exclusion:** Rivals' classroom computer availability is determined by their own budgets, grant allocations, and subsystem rollout decisions made in the prior period at those schools, not by contemporaneous shifts in the school's unobserved appeal. Conditional on track fixed effects and observable school characteristics, the remaining variation in the competitors' equipment reflects rivals' investment timing rather than the school's unobserved heterogeneity. Thus, the instrument shifts the school's chosen value-added only through competitive substitution, not by directly increasing its mean utility.

Overall, our instrument follows the logic of rivals' product characteristics in empirical industrial organization literature: Competitors' resources shift school's quality choice through substitution while remaining orthogonal to the school's unobserved demand shock after controls (Berry et al., 1995). Lastly, we estimate β^q for Equation (4) with Two-Stage Least Squares (2SLS). We control observed school quality with year 2018 information as mentioned in 5.1, including percentage of female teacher, student-teacher ratio, percentage of admitted female students, percentage of admitted students graduated from private middle school, percentage of admitted students graduated from general academic middle school, average middle school GPA among admitted students, indicator on whether school is in Mexico City, number of specialties, and school track fixed effects.

6.3 Estimation of Supply

Our estimation of schools' determination of quality relies on Equation (6), which is derived from the school model. Note that the linear marginal cost on value-added κ enters every term in Equation (6). Therefore, it is not identified. We further normalize κ to 1 so that one unit of cost corresponds to the cost of raising average student value-added by one unit, which is one unit in the ENLACE score. We further make two function specifications. First, we specify the function of schools' value of total enrollment:

$$F_j\left(D_j(\mathbf{q})\right) = \alpha D_j(\mathbf{q}), \tag{10}$$

where α captures the marginal value from each additional student enrollment, which is determined by competition among schools through their value-added choices. Second, we allow the intrinsic value of providing a certain level of value-added differentiated by elite and non-elite schools. Specifically,

$$G_j(q_j) = G_E \cdot \mathbb{1}_{\{\text{elite}_j\}} \cdot q_j + G_N \cdot (1 - \mathbb{1}_{\{\text{elite}_j\}}) \cdot q_j, \tag{11}$$

where $\mathbb{1}_{\{\text{elite}_j\}}$ is an indicator function, equals to one means school j is elite school. Putting the derivative of Equation (10) and Equation (11) into Equation (6), and nomalizing $\kappa = 1$ generate us the final equation for estimation:

$$q_j^* = \alpha D_j' + [(G_N - \eta_N)D_j^{-1} + ((G_E - \eta_E) - (G_N - \eta_N))\mathbb{1}_{\{\text{elite}_j\}} D_j^{-1} - 1]\frac{1}{\sigma_j} - \beta T_j - \omega_j, \tag{12}$$

where q_j^* is the estimated school value-added, σ_j is public school j's own-value-added semi-elasticity of demand, D_j is the demand of school j and D_j' is the marginal effect of value-added on demand. The variable T_j is school j's monthly average teacher wage, which serves as the cost shifter, and ω_j is the measurement error on marginal cost. Lastly, α, $G_E - \eta_E$, $G_N - \eta_N$, and β are parameters to be estimated. $G_N - \eta_N$ represents the net direct incentive to improve quality for non-elite schools, measuring how much they value higher value-added independent of enrollment, after accounting for the fixed cost of providing quality. Analogously, $G_E - \eta_E$ captures the corresponding net direct incentive for elite schools. β measures how changes in schools' teacher wages affect the marginal cost of producing value-added.

Each parameter is identified from distinct sources of cross-school heterogeneity in demand primitives. The coefficient α is identified from variations in the marginal slope of demand with respect to quality. Schools whose enrollment responds more strongly to improvements in quality face larger demand-driven incentives, and this cross-sectional heterogeneity in demand responsiveness pinpoints α.

The coefficient $G_N - \eta_N$ multiplied by $\frac{D_j^{-1}}{\sigma_j}$ highlights settings in which non-demand motives become particularly salient.[14] For example, smaller schools (high D_j^{-1}) and schools facing inelastic demand (low σ_j) load more heavily on this factor. In these environments, demand-driven

[14]Algebraically, $\frac{D_j^{-1}}{\sigma_j} = \frac{1/D_j}{(D_j'/D_j)} = 1/D_j'$. We retain the ratio form to make explicit that identification relies on both school size (D_j^{-1}) and demand responsiveness ($1/\sigma_j$).

incentives are weak: Quality does not meaningfully change enrollment because the baseline scale is small or the semi-elasticity is low. If such schools nonetheless supply high value-added, that residual variation is attributed to non-demand motives. Cross-school heterogeneity in size and elasticity, therefore, provides the identifying variation needed for $G_N - \eta_N$. Finally, the elite premium $(G_E - \eta_E) - (G_N - \eta_N)$ is identified by comparing how $\frac{D_j^{-1}}{\sigma_j}$ correlates with equilibrium quality in the elite sector versus the non-elite sector, and isolates the additional reputational benefits enjoyed by elite schools. In other words, variations in $\frac{D_j^{-1}}{\sigma_j}$ across both elite and non-elite groups yield the leverage to recover the reputational component of elite status, separately from the enrollment-mediated term $\alpha D_j'$.

A central challenge in estimating equation Equation (6) is that both demand D_j and the semi-elasticity σ_j are endogenous to schools' supply choices. Enrollment outcomes are endogenous because they reflect the interaction of two forces: (i) Students choose schools based on their preferences, and (ii) schools adjust quality in response to these choices. As a result, observed demand primitives such as D_j and σ_j are equilibrium outcomes, potentially correlated with unobserved cost shocks that also affect supply decisions. To address this concern, we exploit variations in demand shifters that alter the attractiveness of a school to students but are plausibly orthogonal to schools' marginal cost of raising quality. These demand shifters serve as instruments that shift the demand curve faced by a school without directly affecting its cost structure.

Concretely, we use several geographic and institutional features as demand shifters. First, *weighted distance from students' residences to the school* provides exogenous variations in accessibility that affect enrollment probabilities but not the unobservable school cost. Second, the *presence of nearby public transit stops* lowers travel costs and increases a school's attractiveness without altering its pedagogical costs.[15] Third, the *number of alternative schools in close proximity* captures the intensity of local competition: Schools located in dense markets face lower demand elasticities, while those in thin markets face more captive demand. Finally, *the last year's admission cutoff* proxies for congestion and selectivity shocks that influence perceived desirability but are predetermined with respect to current year cost shocks.

We estimate Equation (6) using the Generalized Method of Moments (GMM). Let Z denote the vector of instruments (demand shifters). The corresponding moment condition is given by $E[\omega_j \mid Z, T_j] = 0$. The instruments include the distance from students' residential postal code to

[15] We obtain public transit stops from Overpass Turbo Website. https://overpass-turbo.eu

school, weighting by number of admitted students, number of COMIPEMS schools within 5km, distance to the nearest COMIPEMS school, distance to the nearest public transit, number of public transit stops within 1km, number of public transit stops within 2km, and 2018 cutoffs. Moreover, we also create interaction terms between the elite indicator with distance to the nearest transit, weighted distance from students' residential postal code to school, and the number of schools within 5km.[16]

7 Estimation Results

This section presents our estimation results. We first present parameter estimates of students' preferences regarding demographics and mean utility. We then present our 2SLS results, which address endogenous value-added. Lastly, we report our estimates from the school model.

7.1 Demand Estimates

Table 4 presents demand estimates of β^x and γ from Equation (3). The estimation results with student demographic controls highlight several systematic patterns in school choice behavior. Students show a strong preference for nearby schools, as indicated by the significant negative effect of distance. A negative and statistically significant coefficient on female suggests that, holding other factors constant, female students obtain lower utility from COMIPEMS schools and are therefore less likely than male students to choose any COMIPEMS school rather than the outside option. Socioeconomic status also shapes choices. Higher household income and maternal education are associated with significantly lower utility for the COMIPEMS alternative, reducing the probability of choosing a COMIPEMS school over the outside option. A plausible explanation is that these households can afford costly schooling options, such as private schooling.[17] In contrast, higher-achieving students, measured by 9th grade GPA, prefer COMIPEMS schools.

Figure 5 displays the kernel densities of estimated mean utility δ_j by school type from Equation (3). The results reveal clear stratification in student preferences: general academic elite schools occupy the highest utility range, with distributions concentrated at the upper tail, reflecting their strong attractiveness. Technological elite schools also show relatively high mean utility, though with greater dispersion. By contrast, general academic non-elite and technolog-

[16] All distances are measured in km.

[17] Note that school fees in the COMPIEMS system are free.

ical non-elite schools exhibit broader distributions that overlap with both low and moderately positive utility values. This suggests that while many non-elite schools are weakly preferred, a subset of them remains highly attractive to certain students, resulting in notable heterogeneity within these categories. Finally, vocational schools cluster tightly around zero, indicating limited overall appeal. Taken together, these patterns underscore both the dominance of elite schools and the persistent variation in student demand across non-elite schools, with some non-elite options emerging as competitive alternatives in the choice set.

Lastly, Table 5 reports 2SLS estimates from Equation (4) using the instrument introduced in Section 6.2. Panel A shows the second-stage results. The table also reports robust standard errors along with their corresponding p-values. In Panel A, the coefficients for rescaled value-added, average middle school GPA of admits, and school in Mexico City are statistically significant at the 1 percent level.[18] The results indicate that students prefer higher-quality schools: mean utility increases significantly with school math value-added. Among covariates, the coefficient on average middle school GPA among admits is strongly positive, and the coefficient on the student–teacher ratio is significantly negative, indicating that students prefer more selective schools and schools with smaller classes. Schools located in Mexico City have substantially lower mean utility, whereas students prefer schools with a higher share of female admits. The coefficients for the share of admits from private middle schools and the share of admits from general middle schools are positive, though not statistically significant. The coefficient of the percentage of female teachers is negative and not significant. Regarding model fit, the specification displays a centered R^2 of 0.712.

Panel B reports the estimates for the first stage, where the rescaled value-added is regressed on the set of covariates and the instrumental variable. The instrument, the competitors' number of computers in classroom use, is strongly correlated with value-added. The coefficient is 0.162 and statistically significant at the 1 percent level. Other variables are also correlated with value added, such as the percentage of female admits, the average middle school GPA of admits, the school in Mexico City, and the number of specialties. Panel C reports standard diagnostic tests for the instrumental variable. The Kleibergen–Paap rk LM test rejects the null of underidentification with a χ^2 statistic of 29.58. The Cragg-Donald Wald F-statistic of 65.29 rejects weak identification, ex-

[18] We rescale math value-added by the national ENLACE standard deviation of 100 points, so coefficients are interpreted per ENLACE point.

ceeding conventional thresholds.[19] The statistic F of 61.62 indicates that the instrument is strongly relevant. The Anderson–Rubin χ^2 test rejects the null that the coefficient on value-added is zero at the 1 percent level with a statistic of 14.05, providing inference robust to weak instruments. Finally, the endogeneity test rejects the null of exogeneity for value-added with a χ^2 statistic of 8.457, supporting the use of the IV approach.

Table 4: Demand Estimates: Student Controls

	Coefficient	Std. Err.
General Controls		
Log Distance	-0.213	0.0004
Female	-0.401	0.021
Socioeconomic Status		
Log Income/person	-0.112	0.011
Mexico City	-0.989	0.021
Mother Years of Education	-0.112	0.003
Student Ability		
9th Grade GPA	0.083	0.013
# Students		172,625

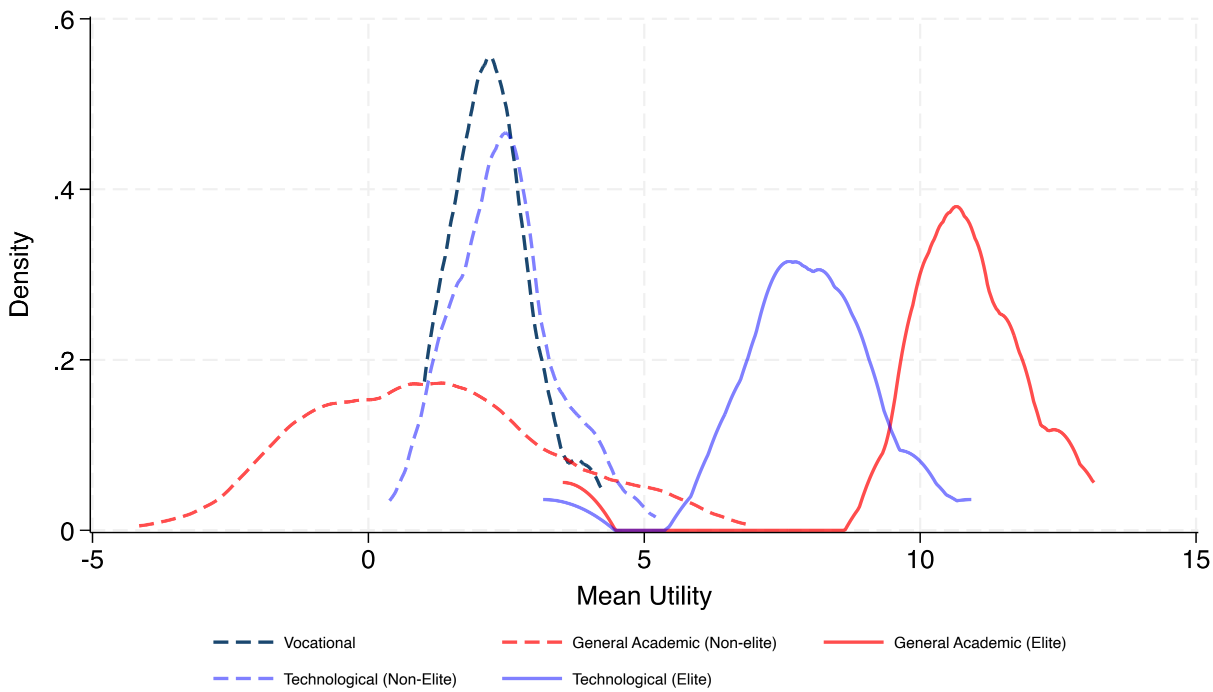
Notes. The table reports estimated preference parameters associated with student characteristics from the demand model under the Stability assumption. The dependent variable is the latent utility of choosing school relative to the outside option. "Coef." gives parameter estimates and "Std. Err." reports robust standard errors. "Log Distance" is the logarithm of the distance between students' residence postal code and the school; "Mexico City" indicates student residence in Mexico City; "Log Income/person" is per-capita household income in logarithms; "Mother's Years of Education" is years of schooling; "9th Grade GPA" measures prior achievement. The last row reports the number of students in the estimation sample.

7.2 Estimates of Suppy Model

We estimate the structural equation in Equation (12) with two-step GMM, where q_j^* is the rescaled value-added in terms of ENLACE points. Table 6 reports the estimation results, where endogenous regressors are identified by demand shifters. The estimated parameter of enrollment value, $\alpha = 0.279$, is positive and statistically significant, indicating that schools derive utility from attracting additional students. Because α is normalized by κ, which is set to 1, its magnitude can be interpreted in units of ENLACE test score value-added: one additional enrolled student raises

[19]The Stock-Yogo critical value for 10% maximal IV size is 16.38.

Figure 5: Distributions of Estimated Mean Utility by School Type



Notes: The figure plots kernel densities of school-level mean utility estimated from the demand model under the Stability Assumption. Each curve shows the distribution of mean utility by school type. Dashed curves represent non-elite tracks, including vocational schools in dark blue, general academic non-elite schools in red, and technological non-elite schools in light blue. Solid curves represent elite tracks, including general academic elite schools in red and technological elite schools in blue. Higher values of mean utility indicate greater average attractiveness in the demand model.

Table 5: IV (2SLS) Estimates

Panel A: Second stage (Dep. var.: *Mean Utility*)			
Variable	Coefficient	Robust Std. Err.	p-value
Rescaled Math Value-added	0.031***	0.000	0.000
% Female Teacher	-0.582	0.517	0.260
Student–teacher Ratio	-0.002*	0.001	0.090
% Female Admits	1.836*	1.063	0.084
% Admits from Private Middle School	4.399	2.703	0.104
% Admits from General Middle School	0.695	0.677	0.305
Average Middle School GPA of Admits	2.910***	0.480	0.000
School in Mexico City	-2.043***	0.195	0.000
# Specialties	0.091	0.106	0.392
Observations		474	
Centered R^2		0.712	
Root MSE		1.401	
Panel B: First stage (Dep. var.: *Rescaled Math Value-added*)			
Variable	Coefficient	Robust Std. Err.	p-value
# Competitor Computer in Classroom	0.162***	0.021	0.000
% Female Admits	-1.493	9.299	0.872
Student–teacher ratio	0.023	0.028	0.412
% Female Admits	-50.731***	14.260	0.000
% Admits from Private Middle School	9.830	38.927	0.801
% Admits from General Middle School	-6.101	10.017	0.543
Average Middle School GPA of Admits	38.782***	6.320	0.000
School in Mexico City	7.076**	3.149	0.025
# Specialties	-4.274**	1.919	0.026
Panel C: Diagnostic tests			
Kleibergen–Paap rk LM χ^2		29.583 (p = 0.0000)	
Kleibergen–Paap rk Wald F		61.616	
Cragg-Donald Wald F		65.286	
Anderson–Rubin χ^2		14.050 (p = 0.0002)	
Endogeneity test χ^2		8.457 (p = 0.0036)	

Notes: In this Table, Panel A reports 2SLS second-stage estimates. The dependent variable is mean utility, and the endogenous regressor is value-added based on math scores, rescaled by the ENLACE standard deviation of 100. The table reports heteroskedasticity-robust standard errors and corresponding p-values. The regression includes school track controls. The unit of observation is the school, and the number of observations is 474. Panel B reports the first-stage estimates for scaled value-added using the excluded instrument, the number of competitors' computers in classrooms use. Panel C reports identification, weak-IV-robust, and endogeneity tests. Statistical significance is denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

the school's objective by roughly 0.28 ENLACE-equivalent points of value-added.

The net direct incentive of non-elite schools, $G_N - \eta_N$, is negative (-0.557) and highly significant, indicating that non-elite schools face a negative direct incentive to improve quality once fixed costs are taken into account. This coefficient is the net direct return to quality: the marginal benefit a non-elite school gains from a one-unit increase in value-added that is unrelated to enrollment, minus the fixed marginal cost of providing that increase. The negative estimate indicates that the fixed cost exceeds the direct non-enrollment benefit, implying that non-elite schools have little or no standalone incentive to raise quality.

For elite schools, the net direct incentive parameter $G_E - \eta_E$ is positive in the point estimate (0.951), but it is not statistically significant. The non-significance is unsurprising given the limited number of elite schools in our sample, which reduces statistical power and precision. The positive coefficient nevertheless suggests that, on average, for elite schools, the direct benefit of increasing quality exceeds the fixed cost. In other words, elite schools derive positive direct returns to raising value-added, such as gains in prestige, reputation, or recognition from educational authorities, in addition to the incentives created by enrollment competition.

The cost sensitivity parameter β is the sensitivity of the cost of improving value-added to changes in wages. In other words, it captures how a school's teacher wage affects the marginal cost of producing value-added. The negative and statistically significant estimate indicates that schools with higher teacher wages face lower marginal costs of improving quality. This pattern is consistent with a selection mechanism in which schools that pay higher teacher wages attract and retain more effective teachers. As a result, these schools face lower marginal costs of improving value-added, since high-quality teachers can raise student performance more efficiently.

We also report instrument validity and fit. The model is overidentified with 11 moments for four parameters. The Hansen J-test fails to reject the overidentifying restrictions $\left(\chi^2(7) = 7.369, p = 0.392\right)$, providing no evidence against the validity of the distance/transit-based instruments and their elite interactions.

Lastly, baseline simulations from the estimated model closely match the observed data (see Appendix D for detailed model fit). This provides confidence in using the model to explore alternative admission schemes in the counterfactual analysis below.

Table 6: Two-step GMM Estimates of Supply Parameters

Parameter	Name	Coefficient	Robust Std. Err.	p-value
Endogenous Regressors				
α	Enrollment value	0.279^{**}	0.133	0.036
$G_N - \eta_N$	Net direct incentive of non-elite schools	-0.557^{***}	0.150	0.000
$G_E - \eta_E$	Net direct incentive of elite schools	0.951	0.940	0.312
Exogenous Regressor				
β	Cost sensitivity	-0.0039^{***}	0.00047	0.000
Observations			474	
# Moments (instruments)			11	
# Parameters			4	
Hansen J (df = 7)		7.369	($p = 0.392$)	

Notes: This table reports two-step GMM estimates for the supply model. The first column lists the parameter symbols; the second column provides plain-language names of the parameters; the third column reports point estimates; the fourth column reports robust standard errors; the last column reports p-values. "Enrollment value (α)" is the marginal value of an additional enrolled student. "Net direct incentive of non-elite schools ($G_N - \eta_N$)" and "Net direct incentive of elite schools ($G_E - \eta_E$)" measure how much non-elite and elite schools, respectively, value higher quality independent of enrollment, after accounting for the fixed cost of providing quality. "Cost sensitivity (β)" captures how teacher wages shift the marginal cost of producing value-added. With the normalization $\kappa = 1$, parameters are expressed in ENLACE value-added units. The number of observations, moments (instruments), parameters, and the Hansen J test are reported below the table. "df" stands for degrees of freedom. Statistical significance is denoted as follows: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

8 Counterfactual Analysis

In this section, we conduct counterfactual simulations under alternative admission rules. We first implement the 2025 scheme while holding school capacities fixed at the 2019 cohort, and report resulting admissions and school quality responses. Motivated by the crowding out effect of the two-list assignment, we then run counterfactuals under the 2025 policy with expanded seats.

8.1 Admission Scheme in 2025

Recall that the 2025 admission policy removes placement test score priorities for non-elite schools and allows each student to submit separate ROLs, one for elite and one for non-elite. For non-elite schools, applicants are admitted directly from their rank lists. When the school is oversubscribed, seats are allocated by a gender-based lottery, with at least 50% of seats reserved for females when feasible, and the remaining seats drawn from the pooled set (males plus unassigned females). Elite schools continue to admit in test score order through the SD mechanism, subject to a middle-school GPA threshold ($\geq$ 7.0).

In our counterfactual, we implement these rules as follows. We simulate student utilities based on the estimated demand model evaluated at the equilibrium school quality. More specifically, using simulated utilities, we construct separate ROLs for each applicant, one ranking non-elite schools and one ranking elite schools.[20] We assume all students submit non-elite lists, and only students with a middle school GPA ≥ 7.0 submit an elite list. Admissions are then processed separately: non-elite seats are allocated by the policy lottery described above, and elite seats are filled by the SD mechanism among GPA-eligible applicants ordered by test score. Thus, a student may receive two offers. Suppose an applicant is placed in both an elite and a non-elite school. In that case, they accept the offer that yields higher utility, and the other seat is released. Note that we do not conduct a complementary round to fill the released seat, given that there are no strict rules on the second-round assignment in the reform. Lastly, schools update their qualities with the admission outcome under the new assignment rules. This implementation mirrors the 2025 rules, allowing us to trace how the admission policy maps into admission, admission composition, and supply quality responsiveness.

Figure 6 reports mean enrollment by school type under the observed baseline rule and the counterfactual. Holding capacities fixed at the 2019 cohort, enrollment falls across all school types once the new assignment rules are placed. The decline is more severe for elite schools than for non-elite (24.75% vs. 12.2%). Enrollment drops by 24.6% in academic elite schools and 24.9% in technological elite schools. The decline in enrollment primarily stems from the design of the new rule. First, a large share of students remain unassigned because, under the two-list system, students can possibly submit both an elite and a non-elite preference list. In particular, all students file a non-elite list, so everyone competes for non-elite seats. In the one-list baseline, students compete for a single pool that includes both elite and non-elite schools, so many high-scoring students already secure seats and do not compete for the remaining seats. By contrast, with the two-list rule, the same number of students now chase a fixed number of non-elite seats, creating congestion. Too many applicants pursue too few spots, and as a result, more students end up without any assignments.

Second, enrollment falls across all school types because, under the two-list rule, a student can temporarily hold both an elite and a non-elite offer. After choosing the option with higher utility, the other seat is released and not reassigned without a second-stage assignment. The substantial decline in enrollment at elite schools indicates that students with two offers are more likely to opt for non-elite schools. This reflects trade-offs between elite and non-elite schools, particularly perceived quality improvements of non-elite schools in the new equilibrium, which will be discussed

[20] Students also rank the outside option for each list.

in the latter part of this section.

Table 7 summarizes the unassigned student composition under the new admission rules. The number of unassigned students rises from 13,815 in the raw data to 36,951 under the counterfactual. Among the unassigned, the share with placement test scores ≥ 80 increases from 9.9% to 15.7%, and the share of students from households below the extreme food poverty line rises from 67.9% to 70.6%. By contrast, the female share among the unassigned declines from 59.4% to 46.4%.

Figure 7 illustrates how the composition of admitted students changes across school types under the counterfactual admission rule compared to the baseline. Panel 7a shows that the share of students with high test scores (≥ 80) increases markedly (on average 10%) in non-elite schools. In contrast, elite schools maintain their high-achieving composition, prioritizing test scores. This pattern is driven by the assignment procedure of non-elite schools: non-elite seats are now allocated by lottery (instead of merit-based), which admits more high scorers into non-elite programs, whereas elite seats continue to be filled by test-score priority, preserving their high-achieving mix. Panel 7b highlights that the proportion of students from households below the extreme food poverty line declines slightly in most school types, suggesting that the counterfactual rule leads to modest socioeconomic sorting away from the disadvantaged students. Panel 7c indicates that gender composition also shifts: non-elite schools admit a higher proportion of female students. The result is consistent with the 50% female quota rule. In oversubscribed schools, at least 50% of the seats are reserved for female applicants, and the remaining seats are allocated from the pooled set, which may still include unassigned female applicants. As a result, the realized female share often exceeds 50%. Taken together, these results suggest that the 2025 admission reform redistributes high-achieving and disadvantaged students across the system, with notable effects on non-elite schools.

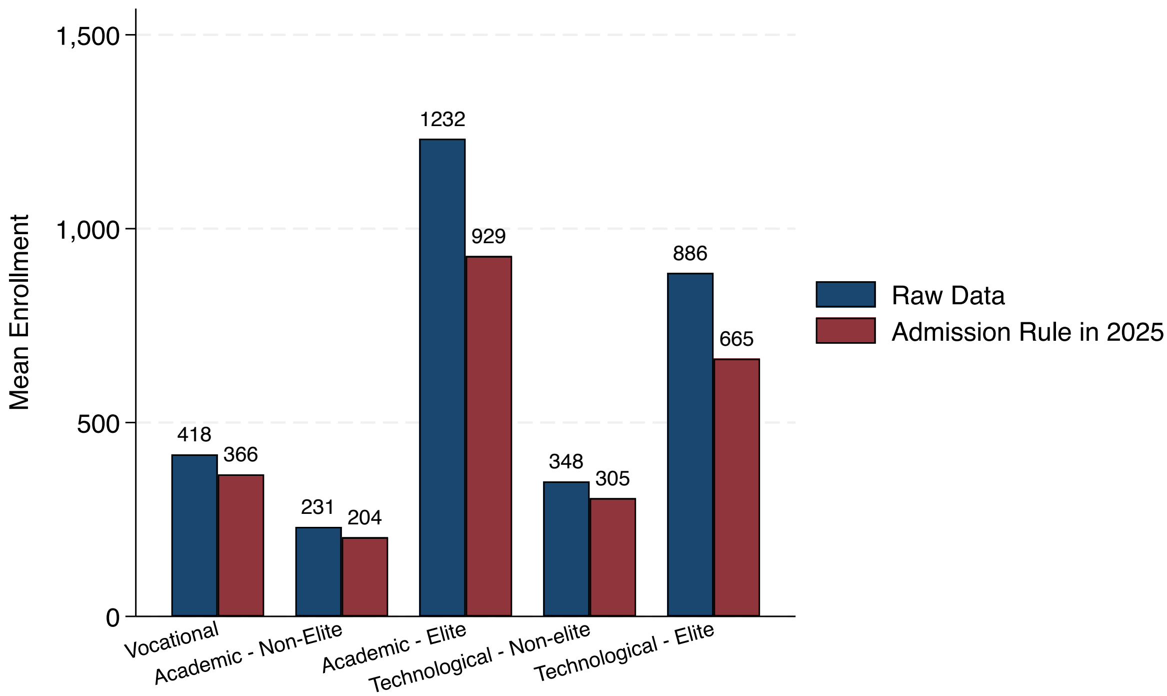
Figure 8 compares the distribution of school quality under the baseline (dashed) and the counterfactual (solid) assignment rules. Panel 8a shows that the value-added distributions of elite schools for both academic and technological tracks shift to the left under the counterfactual. Elite schools experience a sharp decline in enrollment under the two-list rule, and this contraction contributes to a reduction in their quality. As elite schools lose many students, improving quality yields smaller gains but similar costs, making high quality harder to sustain. At the same time, the demand for elite seats becomes less responsive to quality, as eligible students apply regardless of the quality adjustment. With weaker enrollment feedback and higher per-student costs, the marginal return to quality investment declines sharply, leading elite schools to lower their equilibrium quality.

Panel 8b shows that the value-added distribution of non-elite schools by track shifts to the right across all tracks. This improvement arises from a mechanical change in how the assignment rule maps student preferences into realized enrollments. Under the two-list assignment, all students submit ROLs for non-elite schools, prioritizing their first choice. Each non-elite school processes its own pool of first-choice applicants. A slight quality improvement can therefore shift which students rank a given school first, altering the composition of its applicant pool. In contrast, under the one-list SD rule, all schools share a single global priority order based on placement test scores, so small quality differences rarely determine which students are admitted to a given school. As a result, under the 2025 assignment rule, an improvement in non-elite school quality yields a larger change in enrollment probabilities, thereby increasing the effective marginal return to quality.

Lastly, comparing Panel 8a with Panel 8b, although the overall quality dispersion across COMIPEMS schools is reduced in the new equilibrium, we find that elite school quality remains systematically higher than that of non-elite schools under the counterfactual (mean ≈ 50 versus a mean below 10). This reflects a large intrinsic value to providing higher quality among elite schools, even when demand declines.

Overall, the counterfactual analysis suggests that the assignment rule introduced in 2025 has substantially reshaped student admissions and school responses. The two-list design increases the overall number of unassigned students, accompanied by a significant reduction in enrollment at elite schools. The new policy also redistributes high-achieving applicants toward non-elite programs. At the same time, female representation improves in non-elite schools due to the quota rule. Regarding school value-added, the quality dispersion across schools narrows in equilibrium. These results highlight both the equity benefits and the capacity challenges of the reform, demonstrating that future assignment design needs to strike a balance between expanded access and mechanisms to prevent crowding out. For example, policy could include a well-designed supplementary round of assignment or expand access by adding additional seats or incorporting new schools into the system.

Figure 6: Mean Enrollment under Baseline and Counterfactual



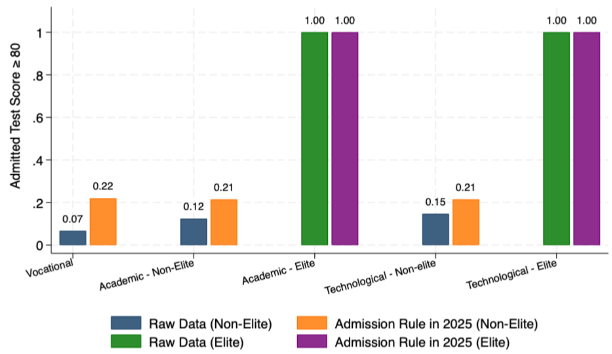
Notes: In this figure, a grouped bar chart shows mean enrollment by school type under two scenarios: the baseline (Raw Data, blue) and the 2025 admission rule (red). The x-axis lists Vocational, Academic Non-Elite, Academic Elite, Technological Non-Elite, and Technological Elite, and the y-axis reports mean enrollment. Numeric labels are displayed above each bar, which are rounded to integers.

Table 7: Unassigned Students under Alternative Admission Rules

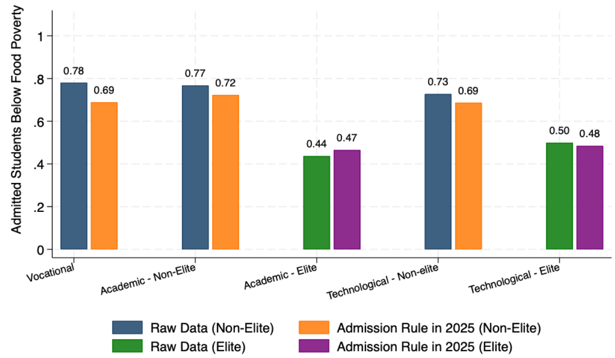
	Counterfactual	Raw Data	Diff.
# Unassigned Students	36,951	13,815	23,100
% Unassigned Students Test Score $\geq$ 80	15.7	9.9	5.8
% Unassigned Students Below Food Poverty Line	70.6	67.9	2.7
% Unassigned Female Students	46.4	59.4	-13.0

Note: The table reports the number of unassigned students and, among the unassigned, the share with placement test scores ≥ 80, below the extreme food poverty line, and female under the counterfactual assignment rules compared with the raw data. “Diff.” is Counterfactual minus Raw Data.

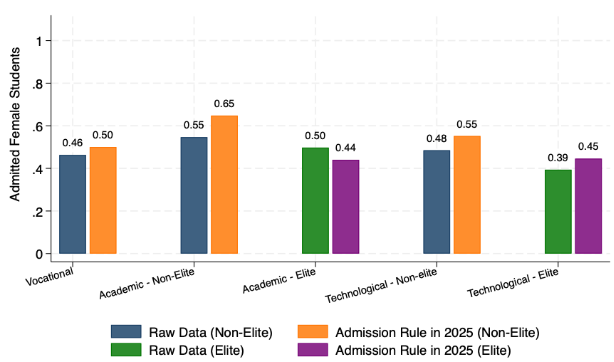
Figure 7: Admission Composition by School Type under Baseline and Counterfactual



(a) Share of Admitted Students with Test Score $\geq$ 80



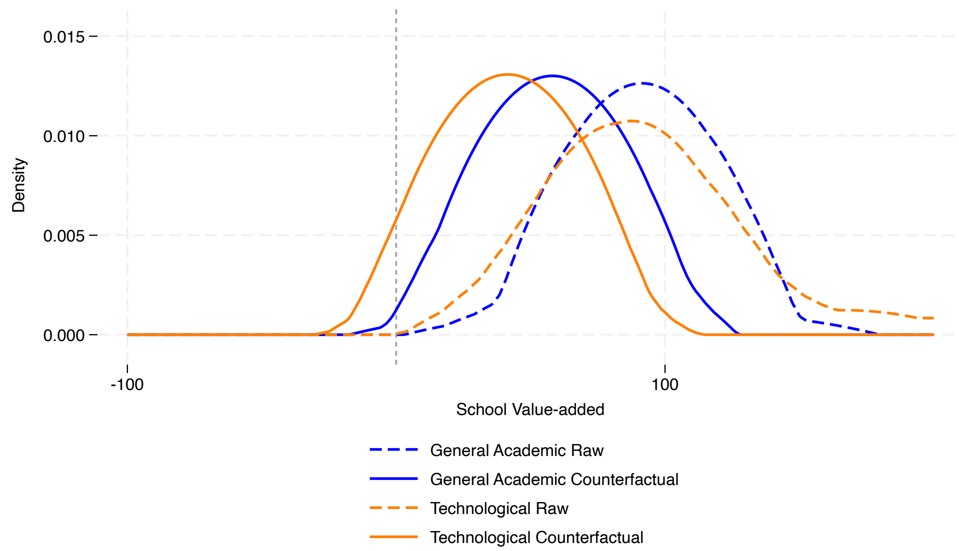
(b) Share of Admitted Students from Household below Extreme Food Poverty Line



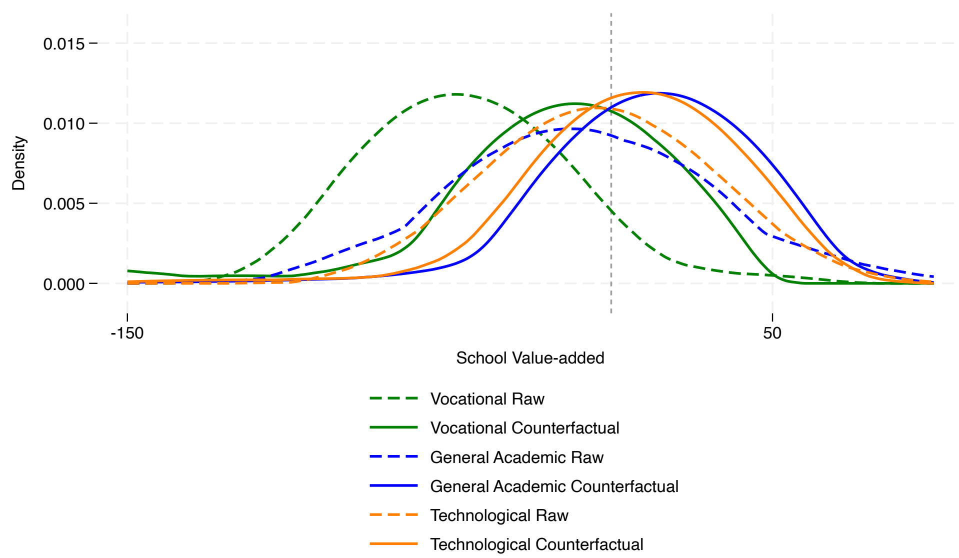
(c) Share of Admitted Female Students

Notes: This grouped bar chart displays the mean admission composition by shares of admitted students by school type under the baseline ("Raw Data") and the counterfactual admission rule. Panel (a) reports the proportion of students admitted with a test score $\geq$ 80, Panel (b) reports the proportion of admitted students from households below the extreme food poverty line, and Panel (c) reports the proportion of admitted female students. Bars are grouped by school type (Vocational, Academic Non-Elite, Academic Elite, Technological Non-Elite, and Technological Elite). Values on top of each bar represent means rounded to two decimal places.

Figure 8: Distribution of School Value-added under Baseline and Counterfactual



(a) Distribution of Elite School Value-added by Track



(b) Distribution of Non-elite School Value-added by Track

Notes: This figure plots the distribution of estimated school value-added under the baseline assignment and the counterfactual assignment rules. Panel (a) shows elite schools by track, and Panel (b) shows non-elite schools, each separated by track. Solid lines represent counterfactual outcomes, and dashed lines represent the observed ("Raw") baseline distributions. School value-added is rescaled by the national standard deviation (100) of ENLACE scores.

8.2 Seat Expansion

As discussed in the previous section, a large number of students remain unassigned under the 2025 assignment scheme, driven by the key feature of the two-list system, while holding capacities fixed at the 2019 cohort. This suggests that limited capacity amplifies the crowding-out effect, leaving more students without placement, which might be the primary concern for the new policy. Therefore, in this section, we run several simulation scenarios in which all non-elite schools expand their seats by 3%, 5%, 7%, and 10%. For schools with fewer than or equal to 10 seats, we allow them to admit three additional students. These expansions correspond to admitting 3,900, 6,439, 8,971, and 12,818 students, respectively.

Table 8 shows how expanding non-elite seats under the 2025 scheme reduces the number of unassigned students, alters enrollment composition, and adjusts school value-added. With capacities fixed at 2019 levels, 36,951 students remain unassigned (Column 2); however, this figure falls steadily as seats expand, dropping to 26,475 when non-elite schools admit 10% more students. Across the seat expansion scenarios, test score composition, the share below the food poverty line, and the female share remain broadly stable. The reason is the random nature of non-elite admissions: additional seats are drawn from the same applicant pool through a lottery (with a 50% female reserve when oversubscribed). At the same time, elite schools continue to admit students in order of test score. As a result, expanding seats mainly reduces the number of unassigned students rather than materially shifting the composition of admitted students. Value-added is also broadly stable across scenarios. Vocational and technological non-elite programs show small increases. Under the 10% expansion, the median vocational school adds approximately 27 seats, and the median technological school adds approximately 34 seats. Elite schools experience a slight decline in value-added when non-elite seats expand, although the magnitude of this effect is minimal.

Overall, the seat expansion does not alter the main patterns observed under the 2025 admission policy; its primary effect is to admit more students. This result highlights seat expansion as a practical lever to soften the unintended side effects of the two-list policy without reshaping its broader distributional impacts.

Table 8: Assignment and Value-added Outcome with Seat Expansion

	Raw Data (1)	Seats in 2019 (2)	Expand 3% Seats (3)	Expand 5% Seats (4)	Expand 7% Seats (5)	Expand 10% Seats (6)
# Unassigned	13,815	36,951	33,756	31,674	29,615	26,475
# Assigned						
Vocational	417	369	379	386	393	403
Academic Non-elite	131	123	126	129	131	136
Academic Elite	886	702	699	696	694	691
Technological Non-elite	324	297	306	312	317	324
Technological Elite	933	690	687	686	684	680
% Assigned Students Test Score $\geq$ ***80***						
Vocational	4.88	20.1	20.0	20.0	19.8	19.8
Academic Non-elite	5.49	20.4	20.3	20.3	20.1	20.1
Academic Elite	1	1	1	1	1	1
Technological Non-elite	11.3	20.7	20.5	20.5	20.4	20.3
Technological Elite	1	1	1	1	1	1
% Assigned Students Below Food Poverty Line						
Vocational	78.2	69.2	69.3	69.3	69.5	69.4
Academic Non-elite	78.1	71.3	71.2	71.2	71.4	71.4
Academic Elite	43.2	46.3	46.3	46.2	46.3	46.2
Technological Non-elite	72.9	68.8	68.9	68.9	68.9	68.9
Technological Elite	50.5	48.5	48.5	48.5	48.3	48.2
% Assigned Female Students						
Vocational	46.6	47.9	47.9	47.7	47.8	47.9
Academic Non-elite	55.3	69.2	68.8	68.8	69.0	68.2
Academic Elite	49.4	44.1	44.1	44.0	44.2	44.5
Technological Non-elite	48.5	50.5	50.4	50.2	50.0	49.9
Technological Elite	35.3	44.8	44.8	44.9	45.0	44.9
Value-added						
Vocational	-46.0	-8.98	-8.65	-8.45	-8.21	-7.95
Academic Non-elite	-7.86	15.06	15.06	14.85	14.99	14.85
Academic Elite	92.29	55.17	55.14	55.14	55.12	55.09
Technological Non-elite	-3.27	11.01	11.06	11.10	11.11	11.37
Technological Elite	87.99	39.37	39.35	39.31	39.29	39.27

Notes: Column (1) reports observed outcomes for the 2019 cohort ("Raw Data"). Column (2) applies the 2025 admission rule while keeping capacities at their 2019 levels ("Seats in 2019"). Columns (3)–(6) show counterfactuals in which non-elite schools expand capacity by 3%, 5%, 7%, and 10%, respectively ("Expand 3%/5%/7%/10% Seats"); elite capacities are unchanged. Small schools with 10 seats are allowed to admit 3 additional students. Counts for "# Unassigned" and "# Assigned" are system totals. All entries are medians across schools within each group (Vocational, Academic Non-elite, Academic Elite, Technological Non-elite, Technological Elite). "Test Score $\geq$ 80" is the share of admitted students with scores $\geq$ 80; "Below Food Poverty Line" is the share of students from households below the food-poverty line; "Female Students" is the female share.

9 Concluding Remarks

This paper examines the consequences of the 2025 admission reform in Mexico's centralized high school assignment system. The reform removed exam-based assignments for non-elite schools and introduced a two-list mechanism that processes elite and non-elite preferences separately. Using administrative microdata and a structural model, we study how the reform reshapes both student allocation and school quality. Our demand estimates indicate that preferences vary strongly with student demographics; elite schools deliver the highest mean utility, while non-elite schools display substantial heterogeneity in perceived quality. The supply-side estimates reveal distinct incentive structures between elite and non-elite schools. The enrollment value parameter is positive and significant, indicating that schools value attracting additional students. For non-elite schools, the negative and significant net direct incentive implies little motivation to raise quality in the absence of enrollment response. For elite schools, they have some direct gains from higher quality, likely tied to reputation. Together, the results suggest that enrollment competition is the primary channel through which quality incentives operate, particularly for non-elite schools.

Counterfactual results suggest that the 2025 admission scheme reallocates a non-trivial share of high achievers to non-elite programs and reduces overall enrollment, assuming capacities remain constant. The enrollment decline is driven by pooling created when students submit separate rank-order lists for elite and non-elite schools under the new rule. Applicants who would have received a single seat under the serial dictatorship now compete on two lists, and more students remain unassigned. Moreover, students with two offers accept one option and release the other seat. Because there is no complementary round, released seats are not reallocated, further lowering filled enrollment. On the supply side, elite schools face weaker demand and consequently lower their quality in equilibrium. In contrast, non-elite schools raise quality due to a stronger sensitivity of demand to perceived improvements. Despite this adjustment, the quality of elite schools remains systematically higher than that of non-elite schools. These results highlight both the equity gains and the capacity challenges of the reform, emphasizing the need for future assignment designs to account for additional factors such as capacity constraints.

We further check the seat expansion scenario to evaluate whether increasing capacity mitigates the crowding out created by the two-list system. Expanding non-elite seats reduces the number of unassigned students substantially while preserving the main qualitative patterns of the reform. The distribution of admitted students and the relative quality adjustments across school types remain similar, suggesting that seat expansion primarily enhances access without altering the broader equilibrium outcomes.

This paper has several limitations. First, our preference estimation relies on feasibility, which is defined by comparing realized exam scores to admission cutoffs. This raises two concerns. First, applicants submit ROLs before the placement exam, so perceived feasibility may be formed with noisy or biased expectations about their own scores; if expected scores differ systematically from realized scores, the inferred choice sets are measured with error, attenuating preference estimates. Second, because we aim to capture school-level competition rather than option (program) level variation, we consolidate multiple options from the same school into a single application. To define feasibility, we use the lowest cutoff among options within the same school as a proxy for access. While this simplifies the choice set, it can overstate availability and reduce the precision of our preference estimates. As a remedy, we plan to adopt the robust stability concept of Artemov et al. (2023), which allows students to have mistaken perceptions about admission probabilities and cutoffs. Under robust stability, preferences are recovered only over clearly feasible schools by adjusting the cutoffs, assuming that students form conservative beliefs about their chances of admission. Moreover, they also conduct tests under alternative cutoff adjustments to assess the precision of preference estimates.

Second, because we do not observe 2025 data, we cannot model the extensive margin of applicants under the new rules. Our counterfactuals hold the applicant pool fixed, so we miss students who would have newly applied (or opted out) when priorities change. A further limitation is that we cannot capture how students might strategically respond to the two-list structure of the new mechanism. For example, in practice, some students may also choose not to participate in the elite exam at all once admission rules separate the two lists. This behavior could arise from perceived low admission chances or application costs. Gender-based strategies could also emerge, as families perceive that female reserves in non-elite lotteries make specific programs easier or harder to access. Without observing actual behavior under the new rules, our model cannot account for these strategic margins. Future work could model how applicants form separate elite and non-elite ROLs, incorporating the decision of whether to take the elite exam, to capture better self-selection, strategic ranking, and beliefs about admission probabilities.

References

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School choice: A mechanism design approach," *American Economic Review*, 2003, *93* (3), 729–747.

_ , Parag A Pathak, and Alvin E Roth, "The New York city high school match," *American Economic Review*, 2005, *95* (2), 364–367.

_ , _ , and _ , "Strategy-proofness versus efficiency in matching with indifferences: Redesigning the NYC high school match," *American Economic Review*, 2009, *99* (5), 1954–1978.

_ , _ , Jonathan Schellenberg, and Christopher R Walters, "Do parents value school effectiveness?," *American Economic Review*, 2020, *110* (5), 1502–1539.

Adelman, Clifford, *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment*, US Department of Education, Office of Educational Research and Improvement, 1999.

Ainsworth, Robert, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola, "Why do households leave school value added on the table? The roles of information and preferences," *American Economic Review*, 2023, *113* (4), 1049–1082.

Akyol, Pelin and Kala Krishna, "Preferences, selection, and value added: A structural approach," *European Economic Review*, 2017, *91*, 89–117.

Allende, Claudia, "Competition under Social Interactions and the Design of Education Policies," Working Paper, Columbia University 2019.

Angrist, Joshua, Peter Hull, and Christopher Walters, "Methods for measuring school effectiveness," *Handbook of the Economics of Education*, 2023, *7*, 1–60.

_ , _ , Parag A Pathak, and Christopher R Walters, "Race and the mismeasure of school quality," *American Economic Review: Insights*, 2024, *6* (1), 20–37.

Artemov, Georgy, Yeon-Koo Che, and YingHua He, "Stable matching with mistaken agents," *Journal of Political Economy: Microeconomics*, 2023, *1* (2), 270–320.

Bau, Natalie, "Estimating an equilibrium model of horizontal competition in education," *Journal of Political Economy*, 2022, *130* (7), 1717–1764.

Berry, Steven T, James A Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Bobba, Matteo, Tim Ederer, Gianmarco Leon-Ciliotta, Christopher Neilson, and Marco G Nieddu, "Teacher compensation and structural inequality: Evidence from centralized teacher school choice in Perú," Technical Report, National Bureau of Economic Research 2021.

_ , Verónica Frisancho, and Marco Pariguana, "Perceived Ability and School Choices: Experimental Evidence and Scale-up Effects," Discussion Paper 16168, Institute of Labor Economics 2023.

Card, David, Martin D Dooley, and A Abigail Payne, "School competition and efficiency with publicly funded Catholic schools," *American Economic Journal: Applied Economics*, 2010, *2* (4), 150–176.

Chetty, Raj, John N Friedman, and Jonah E Rockoff, "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates," *The Quarterly Journal of Economics*, 2014, *129* (1), 1–55.

Combe, Julien, Olivier Tercieux, and Camille Terrier, "The design of teacher assignment: Theory and evidence," *The Review of Economic Studies*, 2022, *89* (6), 3154–3222.

Cullen, Julie Berry, Brian A Jacob, and Steven Levitt, "The effect of school choice on participants: Evidence from randomized lotteries," *Econometrica*, 2006, *74* (5), 1191–1230.

Cunha, Jesse M and Trey Miller, "Measuring value-added in higher education: Possibilities and limitations in the use of administrative data," *Economics of Education Review*, 2014, *42*, 64–77.

de Hoyos, Rafael, Ricardo Estrada, and María José Vargas, "Predicting individual wellbeing through test scores: evidence from a national assessment in Mexico," Policy Research Working Paper, The World Bank 2018.

Delgado, Victor, Jintao Sun, and Rui Zeng, "Spatial Inequality and High School Choice: Implications from Mexico City," Working Paper, Rice University 2025.

Du, Tianyu, Ayush Kanodia, and Susan Athey, "Torch-Choice: A PyTorch Package for Large-Scale Choice Modelling with Python," *arXiv preprint arXiv:2304.01906*, 2023.

Epple, Dennis and Richard E Romano, "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, 1998, *88*, 33–62.

_, Francisco Martinez-Mora, and Richard Romano, "Charter school practices and student selection: An equilibrium analysis," Technical Report, National Bureau of Economic Research 2021.

Fack, Gabrielle, Julien Grenet, and Yinghua He, "Beyond truth-telling: Preference estimation with centralized school choice and college admissions," *American Economic Review*, 2019, *109* (4), 1486–1529.

Figlio, David and Cassandra MD Hart, "Competitive effects of means-tested school vouchers," *American Economic Journal: Applied Economics*, 2014, *6* (1), 133–156.

Figlio, David N, Cassandra MD Hart, and Krzysztof Karbownik, "Effects of maturing private school choice programs on public school students," *American Economic Journal: Economic Policy*, 2023, *15* (4), 255–294.

Gale, David and Lloyd S Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, 1962, *69* (1), 9–15.

Gilraine, Michael, Uros Petronijevic, and John D Singleton, "School choice, competition, and aggregate school quality," Technical Report, National Bureau of Economic Research 2023.

Goodman, Joshua, "The labor of division: Returns to compulsory math coursework," *Journal of Labor Economics*, 2019, *37* (4), 1141–1182.

Haeringer, Guillaume and Flip Klijn, "Constrained school choice," *Journal of Economic Theory*, 2009, *144* (5), 1921–1947.

Hastings, Justine, Thomas J Kane, and Douglas O Staiger, "Heterogeneous preferences and the efficacy of public school choice," Working Paper 12145, National Bureau of Economic Research 2006.

Herrmann, Mariesa, Elias Walsh, and Eric Isenberg, "Shrinkage of value-added estimates and characteristics of students with hard-to-predict achievement levels," *Statistics and Public Policy*, 2016, *3* (1), 1–10.

Hoxby, Caroline M, "Does competition among public schools benefit students and taxpayers?," *American Economic Review*, 2000, *90* (5), 1209–1238.

Imberman, Scott A, "Achievement and behavior in charter schools: Drawing a more complete picture," *The Review of Economics and Statistics*, 2011, *93* (2), 416–435.

Jackson, C Kirabo, Shanette C Porter, John Q Easton, Alyssa Blanchard, and Sebastián Kiguel, "School Effects on Socio-Emotional Development, School-Based Arrests, and Educational Attainment," *American Economic Review: Insights*, 2020, *2* (4), 491–508.

Joensen, Juanna Schrøter and Helena Skyt Nielsen, "Is there a causal effect of high school math on labor market outcomes?," *Journal of Human Resources*, 2009, *44* (1), 171–198.

Kane, Thomas J and Douglas O Staiger, "Estimating teacher impacts on student achievement: An experimental evaluation," Working Paper 14607, National Bureau of Economic Research 2008.

Koedel, Cory and Jonah E Rockoff, "Value-added modeling: A review," *Economics of Education Review*, 2015, *47*, 180–195.

Ladd, Helen F and Randall P Walsh, "Implementing value-added measures of school effectiveness: getting the incentives right," *Economics of Education review*, 2002, *21* (1), 1–17.

Long, Mark C, Patrice Iatarola, and Dylan Conger, "Explaining gaps in readiness for college-level math: The role of high school courses," *Education Finance and Policy*, 2009, *4* (1), 1–33.

Matzkin, Rosa L, "Nonparametric identification and estimation of polychotomous choice models," *Journal of Econometrics*, 1993, *58* (1-2), 137–168.

McMillan, Robert, "Competition, incentives, and public school productivity," *Journal of Public Economics*, 2004, *88* (9-10), 1871–1892.

Meyer, Robert H, "Value-added indicators of school performance: A primer," *Economics of Education Review*, 1997, *16* (3), 283–301.

Navarro, Salvador and Marco Pariguana, "Time Varying Effects of Elite Schools: Evidence from Mexico City," Discussion Paper 316, University of Edinburgh 2025.

Neilson, Christopher A., "Targeted Vouchers, Competition among Schools, and the Academic Achievement of Poor Students," Working Paper, Princeton University 2021.

_ , "The Rise of Coordinated Choice and Assignment Systems in Education Markets Around the World," Working Paper, World Bank 2024.

Ni, Yongmei, "The impact of charter schools on the efficiency of traditional public schools: Evidence from Michigan," *Economics of Education Review*, 2009, *28* (5), 571–584.

Pathak, Parag A and Tayfun Sönmez, "School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation," *American Economic Review*, 2013, *103* (1), 80–106.

Rose, Heather and Julian R Betts, "The effect of high school courses on earnings," *Review of Economics and Statistics*, 2004, *86* (2), 497–513.

Urquiola, Miguel, "Does school choice lead to sorting? Evidence from Tiebout variation," *American Economic Review*, 2005, *95* (4), 1310–1326.

Vivanco, Edgar Franco, "When Change Matters. Identifying Score Gains School Determinants in Mexico: An Intra-cohort Value-added Approach," *Economía Mexicana. Nueva Época*, 2013, *2*, 361–405.

Zimmer, Ron and Richard Buddin, "Is charter school competition in California improving the performance of traditional public schools?," *Public Administration Review*, 2009, *69* (5), 831–845.

Appendix A Details on Assignment Mechanisms

A.1 Serial Dictatorship (SD) Mechanism in COMIPEMS

This appendix section provides the formal definition of the Serial Dictatorship (SD) mechanism, which governs student assignment in COMIPEMS prior to 2025.

The SD mechanism uses students' standardized exam scores to create a strict priority order over all applicants. Each student submits a rank-ordered list (ROL) of programs (institution–campus–track options). The assignment proceeds sequentially: Students are considered one by one, in descending order of exam score, and each is assigned to the highest-ranked program that still has available capacity.

Formally, let students be ordered by priority index $s_1 > s_2 > \cdots > s_N$, where s_i is student i's exam score. Each program j has capacity q_j. The mechanism proceeds as follows:

1. **Step 1.** The student with the highest score, s_1, is assigned to the top-ranked program on her list with $q_j > 0$. Once assigned, that program's capacity is reduced to $q_j \leftarrow q_j - 1$.
2. **Step 2.** The next student, s_2, is assigned to the most-preferred program on her list that still satisfies $q_j > 0$; again, the corresponding capacity is updated to $q_j \leftarrow q_j - 1$.
3. **Step k.** In general, student s_k is assigned to the highest-ranked option on her ROL that has remaining capacity $q_j > 0$ after the first $k - 1$ students have been placed.

When multiple students have identical exam scores, priorities are broken randomly: a lottery is drawn once to induce a complete strict ordering before the procedure begins, and the assignment then follows this order.

The process continues until all students are assigned or all listed options are filled. Because each seat is allocated permanently when a student is considered, no reallocation or "holding and rejection" occurs (in contrast to the deferred acceptance mechanism).[21] In the COMIPEMS context, this implies that elite subsystems (UNAM and IPN) fill their seats early, so access to these options is concentrated among the very top scorers, while subsystems with larger capacity (e.g., SE, CONALEP) absorb much of the remaining demand once selective options have filled.

A.2 Assignment Mechanism for Non-elite Schools in 2025

This appendix section provides the formal definition of the admission mechanism for non-elite schools in 2025, which is an immediate acceptance procedure. Each student submits a rank-ordered list (ROL) of options. The assignment proceeds round by round: in each round, students

[21] In the student-proposing deferred acceptance mechanism, students apply to their most-preferred program in the first round, and programs temporarily "hold" applicants up to capacity while "rejecting" the rest. Rejected students then apply to their next choice in subsequent rounds. This process continues until no further rejections occur. By contrast, under serial dictatorship, once a student is assigned to a program, the seat is never reallocated.

apply to the highest-ranked option on their list that has not yet rejected them, and options allocate seats up to their capacity. Once a student is admitted to an option, that assignment is final and cannot be revoked in subsequent rounds.

Formally, suppose students have ROLs over options. Each option j has capacity q_j. The mechanism proceeds as follows:

1. **Round 1.** Every student applies to her top-ranked option.
   - If the option is *low-demand* ($q_j \geq$ # applicants), all applicants are admitted directly.
   - If the option is *high-demand* (q_j <# applicants), a lottery is conducted to allocate seats. This lottery incorporates a *gender quota*, ensuring that at least 50% of the seats are awarded to female applicants, when necessary.
   - Students who gain admission are permanently assigned to that option. Students not admitted proceed to Round 2.
2. **Round 2.** Each unassigned student applies to her second-ranked option.
   - If the option has remaining capacity, admissions proceed as in Round 1: low-demand options admit all applicants directly, while high-demand options conduct a lottery with a gender quota.
   - If the number of applicants exceeds the remaining capacity, the lottery is restricted to those applicants in the current round and allocates only the available seats, subject to the gender quota.
   - If the option is already full (remaining $q_j = 0$), no lottery is conducted and all applicants are immediately rejected.
3. **Round k.** In general, in round k, each unassigned student applies to the option ranked k on her ROL.
   - If the option has available seats, the same rules apply: low-demand $\rightarrow$ admit all, high-demand $\rightarrow$ lottery with gender quota.
   - If the number of applicants exceeds the remaining capacity, a lottery restricted to the current-round applicants allocates the seats, subject to the gender quota.
   - If the option is already full, no lottery is conducted and the applicant is rejected and proceeds to Round $k + 1$.

The process continues until all students are assigned or all listed programs are filled.

Appendix B More Details on Value-added

B.1 Empirical Bayes Shrinkage

The Empirical Bayes (EB) shrinkage estimator of school value-added is defined as

$$\hat{\theta}_s^{EB} \;=\; a_s \hat{\theta}_s \;+\; (1 - a_s)\, \bar{\theta},$$

where $\hat{\theta}_s^{EB}$ is the shrunken value-added estimate for school s, $\hat{\theta}_s$ is the raw (unshrunken) estimate that we estimated from Equation (1), and $\bar{\theta}$ is the average school effect. The shrinkage weight is given by

$$a_s \;=\; \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{\lambda}_s},$$

where $\hat{\sigma}^2$ is the estimated variance of value-added (after correcting for estimation error) and $\hat{\lambda}_s$ is the estimated error variance of value-added (e.g., the squared standard error). Note that in our data, some schools appear in only one year. In such cases, the school dummy is identical to the year dummy; including both sets of dummies creates perfect collinearity in the design matrix, which prevents standard errors from being computed for all schools. To address this problem, we run a value-added model without year fixed effects:

$$y_{ist} = X'_{ist}\beta + \sum_s 1\left\{S_{ist} = s\right\} \theta_s^f + \varepsilon_{ist}.$$

We use standard error of $\hat{\theta}_s^f$ to calculate empirical shrinkage. More specifically, let the cross-school mean and variance of the raw effects be

$$\bar{\theta} = \frac{1}{S}\sum_{s=1}^{S} \hat{\theta}_s, \quad v^{tot} = \mathbb{V}_s\left(\hat{\theta}_s^f\right).$$

The mean estimation variance across schools is

$$\bar{\lambda} = \frac{1}{S}\sum_{s=1}^{S} \hat{\lambda}_s, \quad \text{where } \hat{\lambda}_s = \left[\widehat{\text{se}}\left(\hat{\theta}_s^f\right)\right]^2.$$

Finally, the between-school variance of true value-added is estimated using the standard EB correction:

$$\hat{\sigma}^2 \;=\; \max\left\{0,\; v^{tot} - \bar{\lambda}\right\}.$$

This ensures that the estimated variance of true school effects is non-negative and accounts for the noise in raw fixed-effect estimates.

B.2 Value-added Estimates

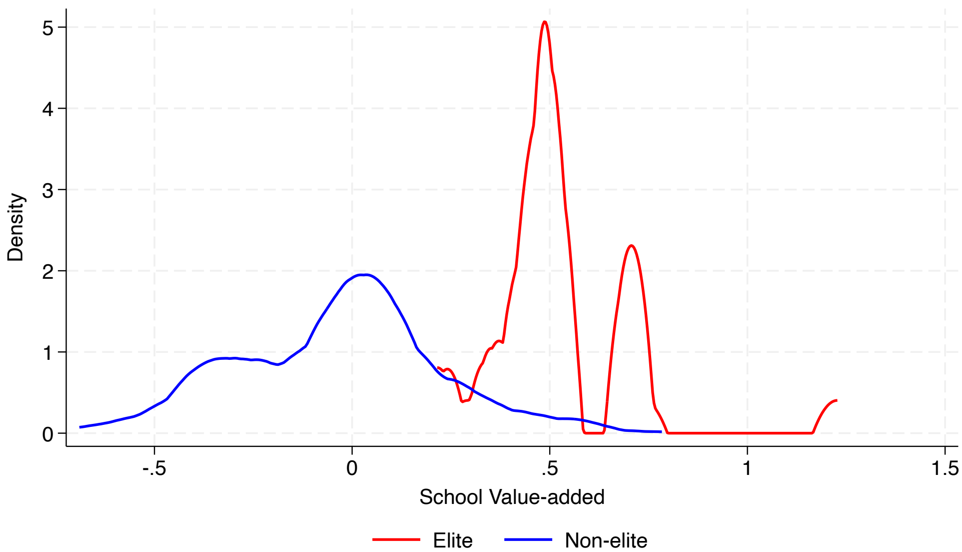
Table B1: Estimated School-level Value-added Models

	Math		Spanish	
	Coefficient	Std. Err.	Coefficient	Std. Err.
Student Demographics				
Female	-0.599***	0.086	0.060	0.071
Age 15 or Younger	-0.222***	0.061	-0.097	0.055
Age 20 or Older	-0.053***	0.011	0.054***	0.010
Not Single	-0.081***	0.019	-0.070***	0.016
Not Single $\times$ Female	-0.072**	0.023	-0.037	0.022
Part-time Working	0.012	0.008	-0.020**	0.007
Part-time Working $\times$ Female	-0.027*	0.011	-0.014	0.009
Elementary School Information				
Rural Public School	-0.035***	0.007	-0.110***	0.007
Private School	-0.040***	0.008	-0.013	0.009
Indigenous School	-0.170***	0.046	-0.228***	0.057
Community School	-0.247***	0.012	-0.338***	0.011
GPA in Elementary School	0.209***	0.006	0.195***	0.006
Female $\times$ GPA in Elementary School	-0.008	0.006	0.002	0.006
Middle School Information				
Private School	0.063***	0.014	0.079***	0.011
Technical School	0.007	0.005	0.005	0.004
Telecommunication	0.058***	0.015	0.002	0.012
GPA in Middle School	-0.023***	0.006	-0.032***	0.005
Female $\times$ GPA in Middle School	0.056***	0.007	0.024***	0.007
Any Merit-based Scholarship	0.013	0.007	-0.013	0.007
Any Need-based Scholarship	0.032***	0.005	0.025***	0.005
High School Information				
To date GPA	0.277***	0.009	0.215***	0.006
Female $\times$ To date GPA	-0.006	0.011	-0.019**	0.007
Mother Education				
Middle School	-0.005	0.006	-0.017***	0.004
High School	-0.039***	0.008	-0.039***	0.006
Technical	-0.007	0.008	0.015*	0.006
College and Above	-0.029**	0.010	-0.007	0.010
Father Education				
Middle School	0.016**	0.005	0.010*	0.004
High School	0.002	0.007	0.016*	0.006
Technical	0.019*	0.008	0.032***	0.007
College and Above	0.020*	0.009	0.047***	0.008
Constant	-3.274***	0.085	-3.095***	0.067
Observations	175,255		175,282	
Adj. R^2	0.431		0.259	

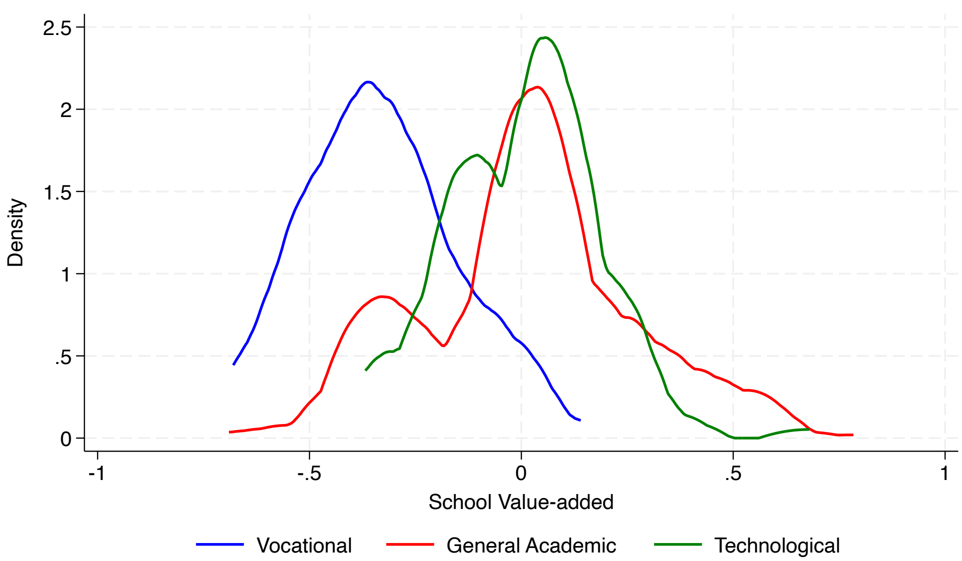
Notes: This table reports the estimation results from the value-added regressions of ENLACE math and Spanish, as represented in Equation (1). Each pair of columns reports a separate regression. The second column presents coefficients for Math, and the third shows the corresponding standard errors. The next two columns report the same for Spanish. School fixed effects and year fixed effects are included; the fixed-effect estimates are not reported. "Coefficient" denotes point estimates and "Std. Err." their standard errors. The set of controls includes student demographics, school characteristics, and performance at the elementary, middle, and high school levels, as well as parental education. Observations and adjusted R^2 are reported at the bottom. Statistical significance is denoted as follows: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B.3 Distribution of Value-added on Spanish

Figure 9: Distribution of Estimated School Value-added on ENLACE Spanish by School Type



(a) Elite versus Non-elite



(b) Non-Elite Schools by Type

Notes: This figure shows the kernel density of estimated school value-added in ENLACE Spanish by school type. Panel (a) compares elite and non-elite schools; Panel (b) decomposes non-elite schools by track (technological, general academic, and vocational).

Appendix C Derivation of Optimal Value-added in School Model

Recall the school model:

$$F_j\left(D_j(\mathbf{q})\right)+G_j\left(q_j\right)-C_j(\mathbf{q}),$$

where $C_j(\mathbf{q})=mc_j\left(q_j\right)D_j(\mathbf{q})+\eta q_j$ and $mc_j\left(q_j\right)=\pi_j+\kappa q_j$, =. Combining the expression of cost functions, rewrite the above function as follows:

$$F_j\left(D_j(\mathbf{q})\right)+G_j\left(q_j\right)-(\beta X_j+\omega_j+\kappa q_j)D_j(\mathbf{q})-\eta q_j.$$

Take the first-order condition with respect to q_j, and set it to 0, we get the expression for setting the optimal quality:

$$F_j'(D_j(\mathbf{q}^*))D_j'(\mathbf{q}^*)+G_j'(q_j^*)-\kappa D_j(\mathbf{q}^*)-(\beta X_j+\omega_j+\kappa q_j^*)D_j'(\mathbf{q}^*)-\eta=0.$$

Re-arranging terms and separating out q^*, we get:

$$\begin{aligned}
q_j^* &= \frac{F_j'(D_j(\mathbf{q}^*))D_j'(\mathbf{q}^*)}{\kappa D_j'(\mathbf{q}^*)}+\frac{G_j'(q_j^*)}{\kappa D_j'(\mathbf{q}^*)}-\frac{\kappa D_j(\mathbf{q}^*)}{\kappa D_j'(\mathbf{q}^*)}-\frac{\beta X_j D_j'(\mathbf{q}^*)}{\kappa D_j'(\mathbf{q}^*)}-\frac{\omega_j D_j'(\mathbf{q}^*)}{\kappa D_j'(\mathbf{q}^*)}-\frac{\eta}{\kappa D_j'(\mathbf{q}^*)}\\
&= \frac{F_j'(D_j(\mathbf{q}^*))}{\kappa}+\frac{G_j'(q_j^*)}{\kappa D_j'(\mathbf{q}^*)}-\frac{D_j(\mathbf{q}^*)}{D_j'(\mathbf{q}^*)}-\frac{\beta X_j}{\kappa}-\frac{\omega_j}{\kappa}-\frac{\eta}{\kappa D_j'(\mathbf{q}^*)}\\
&= \left(\frac{F_j'(D_j(\mathbf{q}^*))}{\kappa}-\frac{\beta X_j}{\kappa}-\frac{\omega_j}{\kappa}\right)+\left(\frac{G_j'(q_j^*)-\eta}{\kappa D_j'(\mathbf{q}^*)}-\frac{D_j(\mathbf{q}^*)}{D_j'(\mathbf{q}^*)}\right)\\
&= \left(\frac{F_j'(D_j(\mathbf{q}^*))}{\kappa}-\frac{\beta X_j}{\kappa}-\frac{\omega_j}{\kappa}\right)-\left(1-\frac{G_j'(q_j^*)-\eta}{\kappa D_j(\mathbf{q}^*)}\right)\frac{D_j(\mathbf{q}^*)}{D_j'(\mathbf{q}^*)}\\
&= \left(\frac{F_j'(D_j(\mathbf{q}^*))}{\kappa}-\frac{\beta X_j}{\kappa}-\frac{\omega_j}{\kappa}\right)-\left(1-\frac{G_j'(q_j^*)-\eta}{\kappa}D_j(\mathbf{q}^*)^{-1}\right)\frac{D_j(\mathbf{q}^*)}{D_j'(\mathbf{q}^*)}.
\end{aligned}$$

Finally expression:

$$q_j^*=\left(\frac{F_j'(D_j(\mathbf{q}^*))}{\kappa}-\frac{\beta X_j}{\kappa}-\frac{\omega_j}{\kappa}\right)-\left(1-\frac{G_j'(q_j^*)-\eta}{\kappa}D_j(\mathbf{q}^*)^{-1}\right)\frac{1}{\sigma_j(\mathbf{q}^*)},$$

where $\sigma_j(\mathbf{q}^*)=\frac{1}{D_j(\mathbf{q}^*)}\frac{\partial D_j(\mathbf{q}^*)}{\partial q_j^*}$ is public school j 's own-value-added semi-elasticity of demand at the optimal level of value-added.

Appendix D Model Fit

We evaluate the model fit for both demand and supply. For the students' choice model, we conduct simulations proposed by Artemov et al. (2023). The simulations account for equilibrium effects induced by changes in program cutoffs, thereby shaping the final matching outcome. In particular, we collect student characteristics, preference estimates, the predicted value of mean utility from 2SLS, and randomly drawn taste shocks from an extreme distribution to simulate students' utilities for schools. According to ordinal utilities, we allow each student to submit a truthful ROL, which differs from the observed ROL in the data. The truth-telling assumption in the simulations is implied by Corollary 2 of Artemov et al. (2023). Intuitively, in the limit economy, the equilibrium strategy results in a unique stable matching, which would be the same as if all applicants submit truthful ROLs, and hence validates the truth-telling assumption when evaluating the matching outcomes in the simulations. We then mimic the SD mechanism with random tie-breaker holding priority indices (placement test scores) and capacities fixed.

We conduct the simulation 500 times and calculate the average admission shares across 500 matching outcomes. Table D2 shows the comparison between actual admission shares and cutoffs observed in the data and the ones calculated from the simulation. In particular, we should look at the shares by demographic composition and program types. The differences are minor for shares among female students and students from households below the extreme food poverty line. Our simulated cutoffs differ slightly from the actual data. This difference arises because, to capture school-level competition, we collapse multiple options/programs into a single choice and use the lowest program cutoff as its threshold. This approach overstates access and lowers the implied cutoffs in the model. Since cutoffs are closely correlated with high test scores, this also helps explain why the simulated share of students with placement test scores of 80 or higher is higher than in the actual data.

Table D3 summarizes how well the estimated school model using GMM matches the targeted moment conditions. With $L = 11$ moments and $K = 4$ parameters, the overidentifying restrictions yield $df = 7$. The Hansen test yields $J = 7.391$ ($p = 0.389$), indicating that we cannot reject the joint validity of the instruments and moment conditions. At the individual level, all standardized residuals are close to zero (absolute values ≤ 1.26), indicating that no single moment condition is statistically different from zero. Some raw residuals (e.g., Weighted Distance from Students' Residential Postal Codes to School) are numerically large, but because their standard errors are also large, the standardized gaps remain small. By contrast, moments with very precise estimates (e.g., Cutoffs in 2018) contribute more to the J statistic despite having tiny raw residuals. Overall, the balanced contributions and modest standardized residuals imply that the model reproduces the selected data moments within sampling error and is consistent with the orthogonality restrictions imposed by the instruments.

Table D2: Model Fit: Admission Outcomes

	Data	Simulated	Diff.
Admission Share by Female			
Elite Academic	0.498	0.465	0.033
Non-elite Academic	0.547	0.486	0.061
Elite Technological	0.394	0.491	-0.097
Non-elite Technological	0.485	0.513	-0.028
Vocational	0.463	0.532	-0.069
Admission Share by Below Extreme Food Poverty Line			
Elite Academic	0.437	0.474	-0.037
Non-elite Academic	0.767	0.648	0.119
Elite Technological	0.499	0.553	-0.054
Non-elite Technological	0.727	0.686	0.042
Vocational	0.780	0.742	0.038
Admission Share by Placement Test ≥ 80			
Elite Academic	1	0.934	0.066
Non-elite Academic	0.124	0.485	-0.361
Elite Technological	1	0.913	0.087
Non-elite Technological	0.147	0.168	-0.022
Vocational	0.067	0.112	-0.046
Cutoffs			
Elite Academic	102	95.973	6.027
Non-elite Academic	38.436	69.568	-31.132
Elite Technological	89.875	81.946	7.929
Non-elite Technological	44.893	53.605	-8.712
Vocational	27.298	43.689	-16.391

Notes: This table reports the model fit on the admission outcome. “Data” reports observed admission outcomes; “Simulated” reports outcomes generated by the model under the 2019 assignment mechanism; “Diff.” equals Data minus Simulated. Each block represents the average percentage of admitted students within each school type. “Placement Test ≥ 80 ” is the share of admits with a placement score of at least 80. “Cutoffs” refer to the type level mean admission cutoff, defined as the placement score of the last admitted student. Type definitions: elite academic, non-elite academic, elite technological, non-elite technological, and vocational follow the text.

Table D3: Model Fit: School-side GMM Moments

Moment	**Moment Residual**	**Moment Std. Err.**	**Standardized Residual**	**Contribution to J**
Weighted Distance from Students' Residential Postal Codes to School	−52.264	192.620	−0.271	1.180
# Schools within 5km	61.460	98.550	0.624	1.554
Distance to the Nearest School	−4.752	11.290	−0.421	−0.831
Distance to the Nearest Public Transit	−5.695	9.523	−0.598	−0.446
# Public Transit Stops within 1km	23.664	19.918	1.188	1.876
# Public Transit Stops within 2km	69.342	64.126	1.081	0.712
Elite × Distance to the Nearest School	−22.016	194.352	−0.113	0.066
Elite × Weighted Distance from Students' Residential Postal Codes to School	−10.906	34.888	−0.313	−0.159
Elite × Distance to the Nearest Public Transit	−4.412	15.521	−0.284	−0.671
Elite × # Schools within 5km	−0.789	1.001	−0.788	1.908
Cutoffs in 2018	−0.714	0.569	−1.255	2.299
Hansen test of overidentifying restrictions: $J = 7.391$, df $= 7$, $p = 0.389$.				

Notes: This table reports the model fit for the GMM moments of the estimation equations in the school model. "Moment Residual" is the empirical moment minus the model-implied moment; positive values indicate the model underpredicts the data. "Moment Std. Err." is the estimated sampling standard error of each empirical moment. "Standardized Residual" reports the moment residual divided by the moment standard error. "Contribution to J" is each moment's contribution to the GMM criterion using the optimal weighting matrix. Distances are in kilometers; counts are numbers of schools or public transit stops within the stated radius; moments labeled "Elite × . . . " are the interaction between the elite indicator and the corresponding variable. The Hansen test of overidentifying restrictions reports the GMM test statistic J, its degrees of freedom (df), and the associated p-value (p).