APPROXIMATING THE EQUILIBRIUM EFFECTS OF INFORMED SCHOOL CHOICE

CLAUDIA ALLENDE
Stanford University

FRANCISCO GALLEGO
PUC-Chile and J-PAL

CHRISTOPHER NEILSON[1]
Yale University, NBER and J-PAL

First version: July 28, 2019.
This version: March 2026.

We study how information about school quality affects school choice and what happens when such an intervention is scaled up. In a randomized trial in Chile, families offered a short video and personalized school report card choose schools with higher test scores and value added. Treated children score 0.2 standard deviations higher on standardized tests, with gains persisting sixteen years later. We embed the experimentally identified demand shift in a structural model of school choice and competition to study equilibrium effects at scale. Capacity constraints and centralized assignment attenuate the demand-side gains, but the same shift in demand strengthens schools' incentives to improve quality, and the supply response more than offsets the congestion effect. Extrapolating directly from the experiment would understate, not overstate, the policy's equilibrium impact. We use the model to guide the design of a follow-up cluster-randomized experiment, focusing on the tradeoff between statistical power and equilibrium spillovers.

KEYWORDS: Information, Incentives, School Choice, Supply Side, Randomized Controlled Trial.

[1]Emails: callende@stanford.edu; fgallego@uc.cl; christopher.neilson@yale.edu. The authors wish to thank Steve Berry, Ryan Cooper, Michael Dinerstein, Francisco Lagos, Chris Walters, and Román Andrés Zárate for useful comments and discussions. We thank the constructive comments from conference participants at the RESTUD Tour 40th Reunion, the NBER Summer Institute IO and Education group meetings and the Y-RISE inaugural conference, as well as seminar and workshop participants at MIT, UCL, UC San Diego, NYU, Princeton, University of Bergen, University of Oslo, PUC-Chile and Yale University. We thank Ryan Cooper for efficiently leading the project development at JPAL-LAC; Josefa Aguirre, Jorge Cariola, José I. Cuesta, Cristián Larroulet, and Cristián Ugarte for excellent research assistance during the fieldwork during 2009-2010; Magdalena Zahri, Ximena Poblete, and Francisca Zegers for help with the production of the information instruments; OPINA and EKHOS for field work; and FONDECYT (Grant No. 1100623) and Industrial Relations Section of Princeton University for financial support.

1. INTRODUCTION

A growing body of evidence shows that providing families with information about school quality can shift their school choices (Hastings and Weinstein, 2008, Andrabi et al., 2017, Kapor et al., 2020, Campos, 2024, Corradini, 2024). But what happens when such a policy is implemented at scale? In a small-scale experiment, school characteristics are, to a first approximation, fixed, and each treated family's choice has negligible effect on market conditions. At scale, this approximation breaks down. Information can shift demand toward the same schools, assignment becomes competitive, and schools' own incentives to invest in quality change. These equilibrium forces—which can amplify or attenuate the partial-equilibrium gains measured in the experiment—are invisible to the experiment itself and cannot be learned without additional economic structure. Whether they matter, and in which direction, is central for deciding whether to implement the policy broadly.[2]

This paper studies these equilibrium effects. We develop a low-cost information intervention, a short video and a personalized school report card, and deliver it to families of preschool children in Chile through existing public infrastructure. A small-scale randomized controlled trial shows that treated families choose schools with higher test scores and value added, especially those reached before they had enrolled. Tracking students with administrative data, we find that treated children score higher on standardized tests up to sixteen years later, including on the college entrance exam. These long-run experimental results establish that the intervention changes behavior in a meaningful way and that the schools families move toward generate better learning outcomes for the students induced to attend them.

We build on the school-choice model in Neilson (2026) to develop a structural framework that can carry the experimentally identified demand shift to the market level.[3] The baseline demand system features heterogeneous families choosing among schools based on quality, out-of-pocket price, and distance across 53 urban markets. Our main addition on the demand side is a treatment-effects parameterization that allows the information intervention both to raise the weight on school characteristics targeted by the report cards and to sharpen choice behavior

[2]The distinction between partial and equilibrium effects has long been recognized as important in education contexts (Heckman et al., 1998). Andrabi et al. (2017) show that information can induce equilibrium responses, including supply-side adjustment, in small village schooling markets in Pakistan. Direct evidence for large urban markets, however, remains limited.

[3]Related work on school choice and competition in education markets includes Neilson (2026), Dinerstein and Smith (2015), Walters (2018), Kapor et al. (2020). The model in this paper also connects to earlier work on school choice in Chile by Gallego and Hernando (2008) and Chumacero et al. (2011).

more broadly, following the scaled-logit logic in Brown and Jeon (2024). These treatment parameters are identified from experimental variation: we estimate them by matching the observed treatment effects on chosen school characteristics to their model-simulated analogs via GMM. On the supply side, we develop a model of school quality choice under capacity constraints and estimate school-specific cost parameters from the quality first-order conditions of privately operated schools. Identification comes from variation in competitive incentives together with cost-of-quality shifters excluded from the demand system, using a ten-year administrative panel across 53 markets. Unconstrained schools respond to demand shifts through enrollment, while capacity-constrained schools respond through student composition, raising quality to attract students eligible for the higher-revenue targeted voucher. The counterfactual equilibrium is then computed as a fixed point: the information treatment makes families more responsive to quality, compressing schools' quality markdowns and inducing quality adjustment that feeds back into demand. We iterate this system to convergence under deferred-acceptance assignment with administrative capacity constraints.

The counterfactual decomposition reveals a clear progression. The decomposition begins with a demand-only benchmark that holds school characteristics fixed and ignores capacity constraints. In that case, the demand shift raises average school quality for low-income families by 0.19 SD at the population level. Adding capacity constraints attenuates this gain by roughly half due to congestion at popular schools. However, the same demand shift also makes families more responsive to quality, reducing schools' quality markdowns and strengthening incentives to improve. In our preferred specifications, this supply-side response more than offsets the attenuation from congestion, so the equilibrium treatment effect exceeds the demand-only benchmark. A partial-equilibrium extrapolation from the experiment would therefore understate, not overstate, the policy's medium-run impact. It would also be silent on whether that effect reflects congestion losses, supply-side gains, or some combination—a distinction that matters for whether the policy improves quality or merely reallocates access to it.

We make three contributions. First, we provide experimental evidence on the effects of an information intervention at the individual level, and use a structural model to study its market-equilibrium effects. Andrabi et al. (2017) provide the closest precedent, showing that school report cards in Pakistan's small village markets raised achievement and induced supply-side responses. In that setting, report cards reach entire communities, so market-level variation arises naturally. Large urban markets, like those we study in Chile, pose a different challenge: schools draw from spatially dispersed populations and families face many alternatives, so demand shifts

diffuse across a wide geography rather than concentrating in a few schools. An experimental design that saturates some urban markets while preserving others as credible control groups may not be feasible, which is why the structural model is needed to study the policy at scale. The counterfactual analysis shows that accounting for these equilibrium forces—congestion and the supply-side quality response—is quantitatively important for the policy assessment.

Second, we contribute to the methodology of combining experiments with structural models for policy evaluation, a literature surveyed in Todd and Wolpin (2023). Our approach lies between two familiar uses of structural models. In one, the policy has already been implemented at scale and the model decomposes observed effects (e.g., Neilson, 2026, Dinerstein and Smith, 2015). In the other, the entire counterfactual rests on the model, with no experimental variation to anchor the demand response. Here, the experiment identifies a credible demand shift that is observed in the data, but the full policy environment—congestion, assignment constraints, supply-side responses—must be studied through the model, because these forces emerge only when the treated population is large enough to change market conditions. A comparison with Neilson (2026) makes the distinction concrete: that paper studies Chile's SEP voucher reform after it was implemented at scale, and uses the model to recover demand and decompose the observed equilibrium response into its underlying competitive forces. For the information policy studied here, no such implementation exists, so the model must recover cost parameters, introduce capacity constraints, and solve for the counterfactual quality equilibrium. In this sense the model is used for prospective extrapolation rather than retrospective decomposition, and the demand shift rests on the experiment rather than on the model alone.[4]

What makes this extrapolation credible is that the structural model and its key assumptions are well grounded empirically. In our setting, families choose among differentiated schools with observable quality measures, prices, and locations. The institutional environment—including voucher subsidies and regulated copayments—is concrete and well documented. We also observe the adoption of centralized assignment after our estimation period, and analyzing those data helps us understand how assignment frictions, congestion, and capacity constraints operate in this market, and thus how they should be incorporated into the policy extrapolation. Administrative data covering enrollment, school finances, teacher characteristics, and student outcomes over a decade provide the variation needed to discipline both demand and supply. These features

[4]Todd and Wolpin (2023) also discuss cases where experiments estimate or validate structural parameters (Todd and Wolpin, 2006, Walters, 2018). A related approach to ours appears in Lise et al. (2015), who combine experimental variation with equilibrium considerations in the context of employment.

also make it possible to subject the equilibrium predictions to transparent robustness checks—varying capacity assumptions, allowing public schools to respond, scaling up input costs, and targeting different subgroups—thereby stress-testing the extrapolation.

Third, we show how the estimated model can inform the design of a larger-scale experiment aimed at detecting equilibrium effects directly. One middle ground between randomizing entire markets and randomizing individual families is to define smaller geographic clusters within markets and assign them independently to treatment or control. The challenge, however, is that when a policy generates equilibrium responses that propagate through shared schools, treated and control clusters are not independent, and the trade-off between cluster size and statistical power depends on the spatial reach of these spillovers. In Section 7, we use the model to navigate this trade-off, studying how cluster size, buffer zones, and take-up rates shape the bias and precision of a cluster-randomized trial in Santiago. The analysis reveals that the optimal design depends on which equilibrium channels are active: buffer zones that are irrelevant under demand effects alone become essential once supply-side quality adjustments create cross-boundary contamination. More generally, the same framework that extrapolates a small-scale experiment to the market level can also guide the design of a larger experiment that could test the equilibrium predictions.

The rest of the paper is organized as follows. Section 2 provides a brief conceptual overview of the economic forces that motivate the analysis. Section 3 describes the institutional setting in Chile, the design of the information intervention, and the data. Section 4 presents the experimental results. Section 5 develops the structural model and reports parameter estimates. Section 6 simulates the equilibrium effects of scaling up the information policy. Section 7 uses the model to design a larger-scale experiment. Section 8 concludes.

2. CONCEPTUAL FRAMEWORK

2.1. *Framework for Policy Analysis*

This section summarizes the policy problem that motivates the empirical analysis. The randomized controlled trial identifies how the intervention changes school choices for the families we study. A policymaker, however, is interested in a different question: what would happen if the same intervention were implemented more broadly, in places where families face different local school environments and where the intervention itself may change market conditions? The experiment identifies the demand shift induced by the policy; the challenge for the rest of the paper is to understand how that shift propagates once the policy is scaled up. This is the sense

in which the paper combines experimental and structural evidence: the experiment provides the demand shift, and the model provides the equilibrium analysis that the experiment alone cannot deliver.

Two features of this setting make that extrapolation nontrivial. First, school choice is highly localized. Families differ in socioeconomic characteristics, voucher eligibility, and residential location, and these differences imply different effective choice environments: the schools that are nearby, their quality, their prices, and the distances required to attend them all vary across families. Even within the same urban market, families living in different locations can face substantially different choice sets. While experimental heterogeneity analysis by observable characteristics—such as socioeconomic status or education level—can reveal which types of families respond more to the treatment, it cannot capture the full richness of how responses vary with the local school environment. A family's effective choice set depends on the number, quality, prices, and distances of nearby schools—a multidimensional spatial object that is difficult to summarize in a single dimension amenable to subgroup analysis. We observe these local configurations in the data, but their complexity is precisely why the structural model is needed to transport the experimental treatment effect across different choice environments. The treatment effect identified in the experimental sample therefore need not transport mechanically to a broader population with a different spatial and socioeconomic composition.

Second, scaling the intervention can generate equilibrium effects. When only a small group of families is treated, school quality, assignment probabilities, and school incentives are, to a first approximation, fixed. When many families are treated, that approximation becomes less credible. Increased demand for higher-quality schools can create congestion, especially when access is mediated by assignment and school capacity. In the medium run, changes in demand can also alter schools' incentives to improve quality. The aggregate effect of the policy may therefore differ from the partial-equilibrium effect measured in a small-scale experiment.

The empirical strategy in the rest of the paper is designed to address these two issues jointly. We combine the experimental variation with a model of school choice that allows treatment effects to vary with local choice environments, then embed that demand system in counterfactual exercises with assignment, finite capacity, and school quality responses. This provides a disciplined way to move from the experimentally identified effect for the treated sample to policy-relevant predictions for larger populations. In particular, it lets us study three objects that a small-scale experiment alone cannot recover: how the policy transports across different local

choice environments, how congestion alters assignments when many families are treated, and how schools adjust quality in response to the resulting competitive pressure.

Beyond these policy counterfactuals, the framework also informs the design of larger-scale experiments. Because the model quantifies how demand shifts propagate spatially and how supply-side responses spill across geographic boundaries, it provides the ingredients needed to evaluate candidate cluster-randomized designs before they are fielded. We apply this logic in Section 7, where we design a trial in Santiago and show that the optimal configuration depends on which equilibrium channels are active.

3. SETTING, INTERVENTION, AND DATA

3.1. *Institutional Setting*

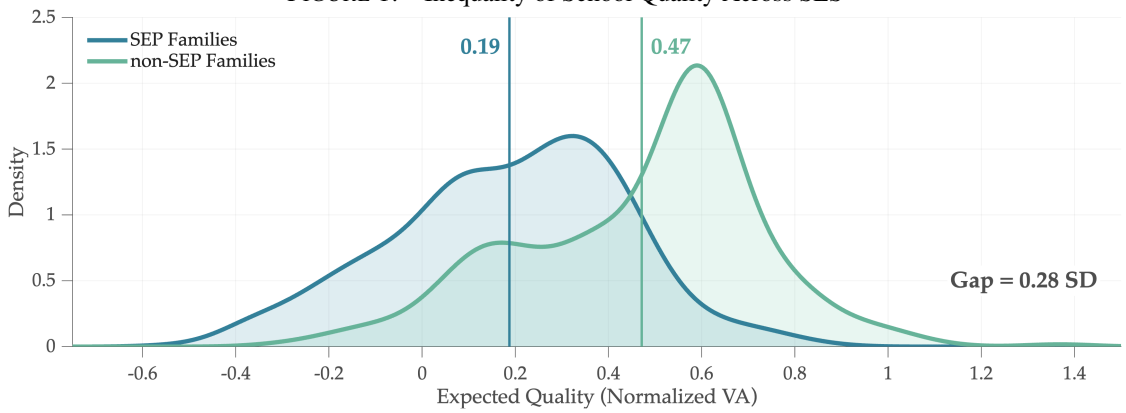
Primary schools in Chile are either free public schools, private voucher schools, or private non-voucher schools. The system features a high degree of school choice and a large private sector: in 2016, the market share for first-grade students was 36% for public schools, 55% for voucher schools, and 8% for private schools. Voucher schools may charge copayments, but these are capped and interact with targeted subsidies for disadvantaged students. In particular, in 2008 the government introduced the SEP reform (*Subvención Escolar Preferencial*), which increased the voucher for the poorest 40% of students and required participating schools to charge zero tuition to eligible families. Subsequent reforms further reduced fees and replaced the earlier decentralized admissions process with a centralized application system.[5]

In spite of the variety of schools and choices available, students from poorer families tend to attend schools with lower academic outcomes and lower inputs. Figure 1 shows the distribution of estimated school value added for students entering first grade by SEP eligibility, which proxies for socioeconomic status. One reason for this gap is that poorer families may place less weight on academic quality when choosing schools, either because quality is harder to observe or because the returns to choosing a better school are less salient. Chilean policymakers have long published school performance data, but this information is rarely tailored to the schools in a family's immediate vicinity or timed to reach parents when they are actively choosing where to enroll. Our intervention targets both gaps.

We reach families through the public preschool system. Pre-K attendance is widespread before primary school entry, with net enrollment rates of 55% at age 3 and 87% at age 4 (OECD,

[5] See Neilson (2026) for a detailed description of the market structure and policy environment.

FIGURE 1.—Inequality of School Quality Across SES



Notes: Distribution of expected school quality (normalized VA) by SEP eligibility. SEP families are eligible for Chile's targeted voucher (*Subvención Escolar Preferencial*), which proxies for low socioeconomic status. Vertical lines mark group means.

2016). We partnered with *Fundación Integra*, a national network of approximately 1,000 public preschool centers serving low-income families throughout the country. We delivered the intervention at preschool parent meetings, reaching families during the period when they are actively deciding which primary school to enroll their child in for the following year. At the time of our experiment (2010), each school managed its own admissions and timing varied across schools. Chile has since adopted a centralized application system with a common enrollment window, which would make the timing of information provision more predictable in a scaled-up policy environment. The Online Appendix provides additional institutional detail on preschool provision.

3.2. *Information Intervention and Experimental Design*

The intervention was designed to make school choice more salient and to provide parents with simple, comparable information about nearby schools. Its goal was to increase both awareness of local school characteristics and the perceived returns to school quality. The treatment was delivered at regular *Integra* parent-teacher meetings and combined a short video with a personalized report card. The video emphasized that choosing a school carefully can shape a child's future opportunities: it linked school quality to preparation for higher education and better jobs, and featured role models from low-income neighborhoods who had accessed better opportunities through education (Figure O-2 and Figure O-3). Each family also received a report card listing test scores, prices, and locations of nearby primary schools, together with a map to help parents locate and compare options (Figure O-1).

We implemented the study in the three largest regions of Chile: Valparaíso, Biobío, and Santiago. Preschools were eligible if they were in urban areas, had at least 10 primary schools within 2 km, and were located in municipalities with at least two primary schools per preschool. We randomly assigned preschools to treatment and control groups, stratifying by region, the number of grades offered, and a measure of local competition given by the number of primary schools within 2 km. The original design distinguished between two treatment intensities, but imperfect tracking during implementation prevents us from separating the two arms. We therefore pool both versions and interpret the estimates as intent-to-treat effects of assignment to the bundled information package. Additional details on sample selection, treatment materials, and implementation appear in Online Appendix Sections O-2 and O-3. Figure O-5 maps the spatial distribution of treatment and control preschools and student homes in Santiago, and Figure O-6 shows how treated students are distributed across destination schools.

The experiment was implemented between August and December 2010 by trained staff who attended preschool parent meetings. In the 133 preschools that agreed to participate, 1,832 parents signed informed consent forms and completed a baseline survey before any information was provided. The baseline survey collected contact information and information on the application process, including whether the parent expected the child to enter primary school in 2011, whether a school had already been chosen, and whether the family had older children already attending primary school. Parents in treatment preschools then received the intervention. Because our presence at the meeting was not announced in advance, parents could not sort into attendance based on interest in school enrollment or demand for information. Between May and July 2011 we conducted a follow-up survey on enrollment decisions and successfully re-interviewed 1,612 of the 1,832 baseline respondents (88%). We also matched 1,795 of the 1,832 parents in the original sample to administrative records using individual student identifiers (98%).

3.3. *Data, Outcomes, and Analysis Samples*

We use several complementary data sources. First, we use administrative data on preschools from *Integra*, including location, enrollment, attendance, and socioeconomic composition. Second, we use self-collected data from the baseline and follow-up surveys in the experimental preschools. These data include contact information, individual identifiers, family location, and questions regarding the application process. We match this information to administrative data from Chile's Ministry of Education.

The Ministry of Education data include student-by-year matriculation records, information on grade progression and demographics, eligibility for the SEP voucher, and standardized test scores from SIMCE. We complement these data with birth records from the Ministry of Health, which contain information on child health at birth and on parental background, and with administrative records on all schools in the country, including school type, enrollment, address, tuition, and other characteristics. For the supply analysis, we also use a school-year panel with transfers, expenditures, and measures of teacher and principal quality.[6]

Different parts of the paper use different slices of these data. The reduced-form analysis uses the experimental cohort treated in 2010 and their observed first-grade choices in 2011. The baseline demand parameters that anchor the structural model are inherited from Neilson (2026), who estimates demand using administrative first-grade choice data from 2006, 2007, 2011, and 2012 across 53 urban markets. Our second-step treatment estimation uses the subset of experimental families who had not yet enrolled and can be matched to the six urban markets shared with the baseline demand model. The supply estimation uses the same 53-market urban sample extended to a 2006–2015 school-year panel, and the main counterfactuals are simulated in the 2011 market environment, the year in which the experimental cohort entered first grade. These sample differences are modest and arise from transparent restrictions such as market overlap, enrollment timing, and the availability of the variables required for each exercise. Online Appendix O-13.1 summarizes how these samples map into each empirical exercise, and Table O-7 reports summary statistics for the main estimation variables.

Our main short-run outcomes are the characteristics of the schools families choose, including distance, fees, test scores, and value added. Our longer-run outcomes are student academic outcomes measured in administrative test-score and attainment records. Because the intervention was delivered during the period in which parents were enrolling children in primary school, timing matters for interpretation. Families that had already chosen or enrolled before the preschool meeting had less scope to respond to the treatment. We therefore report results for both the full sample and the subgroup of families that had not yet matriculated at the time of treatment, which is also the sample used for the structural treatment-effect estimation. We find no meaningful baseline differences between treatment and control preschools or families, and treatment-control balance is preserved within the enrolled and not-yet-enrolled subsamples. Detailed balance tables, additional evidence on enrollment timing, and further discussion of treat-

[6]For further details, see the Online Appendix and Neilson (2026).

ment implementation are reported in Table O-4, Table O-5, Table O-6, and Online Appendix Section O-4.

We use the school-level test-score value-added measure developed in Neilson (2026) as our main proxy for academic quality, extending its coverage as needed for the empirical exercises in this paper. This measure captures the school's contribution to later student achievement after conditioning on a rich set of student and school characteristics. We show in the Online Appendix that it is strongly correlated with teacher quality—in particular, with teachers' own scores on Chile's national university entrance exam (PSU, now PAES)—and that rankings are similar whether or not prior scores are available as controls. The report cards given to treated families displayed raw test scores (SIMCE averages), not value added, since raw scores are publicly available and easy for parents to interpret. The two measures are strongly correlated across schools (Figure O-7), so the quality information provided to parents is closely aligned with the measure used in the structural model. The estimation equation and validation exercises are presented in Online Appendix Section O-7.

4. EXPERIMENTAL RESULTS

Table I reports the effects of the treatment on the characteristics of the schools families chose, and Table II reports effects on student outcomes at longer horizons. All specifications include randomization controls and baseline family controls, and the Online Appendix shows that results are very similar without family controls. Because the intervention was delivered while families were in the process of choosing a primary school, we report results for the full sample and separately by retrospective enrollment status at the moment of intervention. Enrollment status comes from the follow-up survey, which covers 1,612 of the original 1,832 parents, so the pooled specification is estimated on a slightly different sample than the two enrollment-status subsamples.

The experimental evidence on school choice is concentrated among families reached before they had committed to a school (Panel C of Table I). In this group, treated families travel about 0.2 km farther, are 12.5 percentage points more likely to choose a fee-charging school, and choose schools with stronger academic profiles.[7] The strongest shifts are in second-grade language, fourth-grade math, and value added, with a positive but less precise estimate for fourth-

[7]For the distance outcome, we exclude families that appear to have moved, identified by either a different municipality in administrative records or a home-to-school distance exceeding 4 km. Results are robust to alternative distance restrictions.

TABLE I

EFFECTS ON CHARACTERISTICS OF SCHOOLS CHOSEN ONE YEAR AFTER TREATMENT

	Distance (kms)	Positive Price	Lang 2nd	Lang 4th	Math 4th	V. Added[a]
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full Sample						
Treatment	0.077	0.050	0.006	0.010	0.020	0.031
	(0.061)	(0.035)	(0.022)	(0.025)	(0.027)	(0.031)
Control mean	1.081	0.494	-0.119	-0.095	-0.082	-0.024
N obs.	1,427	1,701	1,693	1,692	1,692	1,701
Panel B: Already enrolled at the time of the PreK visit						
Treatment	-0.153	-0.003	-0.026	-0.022	-0.042	-0.030
	(0.130)	(0.055)	(0.043)	(0.048)	(0.058)	(0.060)
Control mean	1.192	0.589	-0.031	-0.007	0.046	0.067
N obs.	512	578	576	577	577	578
Panel C: Not enrolled at the time of the PreK visit						
Treatment	0.200***	0.125***	0.049*	0.041	0.080**	0.095***
	(0.070)	(0.041)	(0.027)	(0.029)	(0.032)	(0.037)
Control mean	1.010	0.434	-0.179	-0.137	-0.146	-0.067
N obs.	800	934	928	926	926	934

Notes: ITT estimates from regressions of chosen-school characteristics on treatment assignment. Panel A: full sample; Panel B: families already enrolled at the time of treatment; Panel C: families not yet enrolled. Standard errors in parentheses are clustered at the preschool-center level. Exact controls and sample restrictions are reported in Appendix A1. [a] Column 6 uses the 2011 school value-added measure Q_2 from Neilson (2026). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

grade language. The value-added effect—about 0.10 in the normalized quality measure used in the model—is the object most directly relevant for the structural analysis. Among families who had already enrolled (Panel B), the point estimates are uniformly small and statistically insignificant, consistent with the treatment having limited scope to change a decision that was already made. In the full sample (Panel A), the estimates are attenuated accordingly.

A potential concern is that the not-yet-enrolled subsample is selected. However, enrollment status was determined before treatment—it reflects when our team visited the preschool center, not differential family behavior—and baseline characteristics are balanced within both subsamples (Table O-6). This is also the most policy-relevant group: under Chile's current centralized

admissions system, families face a common application window, so a scaled-up intervention would reach most families before enrollment. The structural analysis below focuses on this group.

Table II examines whether the changes in school choice translated into better academic outcomes. Among not-yet-enrolled families, fourth-grade SIMCE scores are approximately 0.22 SD higher in language and 0.22 SD higher in math. By eighth grade, the language estimate is smaller and statistically insignificant (0.13 SD), while the math effect remains large and significant (0.27 SD). The treatment does not appear to affect educational attainment: neither high school graduation nor participation in the college entrance exam moves. This pattern is consistent with the intervention changing the quality of schooling rather than the quantity. On the college entrance exam itself (PAES), treated students in the not-yet-enrolled group score 0.18–0.22 SD higher, depending on the subject—magnitudes close to the fourth-grade effects, suggesting fairly persistent gains over time. For the already-enrolled group, the PAES estimates are also positive, though less stable across specifications. These families could not easily switch schools in the short term in response to the treatment, but they nonetheless received information emphasizing the importance of school quality. If the intervention affected educational investments beyond immediate school choice—for instance, by encouraging families to seek better options at later transition points or to engage more actively with their children's learning—effects could emerge over longer horizons even without an immediate school switch. We view these estimates as suggestive and focus the structural analysis on the not-yet-enrolled group, where the treatment had direct scope to affect school choice in the short term.

Finally, we examine whether the report card shifted choices through superficial design features rather than through the information it conveyed. As reported in the Online Appendix, treated families are not more likely to choose schools merely because those schools were listed on the card, color-coded, or highlighted among the top five. If anything, the negative coefficient on choosing a listed school is consistent with treated families looking beyond the card to find better options farther away. These patterns suggest that the intervention worked by increasing the salience of school quality and stimulating search, rather than by mechanically directing families toward particular schools on the card.

In sum, the experimental evidence shows that providing families with personalized information about nearby schools shifts choices toward higher-quality options and produces lasting gains in academic achievement. The effects are concentrated among families reached before enrollment, which is the policy-relevant margin for a scaled-up intervention. These experimental

TABLE II

EFFECTS ON LONG-TERM STUDENT OUTCOMES

	SIMCE Test Scores					College Entrance Exam (PAES)		
	4th Grade		8th Grade		Grad.	Took	(Cond. on taking PAES)	
	Lang	Math	Lang	Math	HS	PAES	Lang	Math
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full Sample								
Treatment	0.087	0.115**	-0.006	0.236**	-0.012	0.023	0.191***	0.174***
	(0.054)	(0.053)	(0.082)	(0.094)	(0.032)	(0.032)	(0.071)	(0.053)
Control mean	-0.222	-0.145	-0.059	-0.103	0.728	0.592	-0.367	-0.225
N obs.	1,583	1,584	610	615	1,775	1,775	1,083	1,060
Panel B: Already enrolled at the time of the PreK visit								
Treatment	-0.085	-0.090	-0.221	0.367**	-0.039	0.007	0.242**	0.292***
	(0.115)	(0.096)	(0.239)	(0.161)	(0.053)	(0.059)	(0.111)	(0.096)
Control mean	-0.039	0.058	0.169	0.044	0.770	0.613	-0.287	-0.145
N obs.	543	533	241	243	596	596	373	364
Panel C: Not enrolled at the time of the PreK visit								
Treatment	0.234***	0.215***	0.125	0.267**	-0.003	0.025	0.217***	0.176***
	(0.076)	(0.065)	(0.111)	(0.108)	(0.039)	(0.033)	(0.080)	(0.067)
Control mean	-0.339	-0.252	-0.188	-0.162	0.706	0.579	-0.447	-0.284
N obs.	857	865	301	301	975	975	581	568

Notes: ITT estimates from regressions of student outcomes on treatment assignment. Panel A: full sample; Panel B: families already enrolled at the time of treatment; Panel C: families not yet enrolled. Columns 1–2: 4th-grade SIMCE; Columns 3–4: 8th-grade SIMCE; Columns 5–6: attainment (high school graduation, college entrance exam participation); Columns 7–8: college entrance exam scores, conditional on taking the exam. Controls and sample restrictions follow Table I. Standard errors in parentheses are clustered at the preschool-center level. Exact specifications are reported in Appendix A1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

patterns—especially the treatment-induced shifts in chosen school quality, price, and distance—provide the empirical targets for the structural analysis in the next section.

5. EMPIRICAL MODEL FOR POLICY ANALYSIS

This section develops the structural model. The baseline demand system—utility specification, market shares, and family types—is inherited from Neilson (2026) and summarized briefly in subsection 5.1. Our main contribution on the demand side is a treatment-effects parameterization (subsection 5.2) that allows the information intervention to shift school choice through characteristic-specific channels and a common rescaling of choice behavior. A Bayesian signal model (subsection 5.3) provides one interpretation of the scale parameter, following Brown and

Jeon (2024). On the supply side (subsection 5.4), we recover cost parameters from schools' quality first-order conditions and introduce a capacity extension that modifies the quality response when schools cannot easily expand enrollment. subsection 5.5 describes the three-step estimation.

5.1. *Baseline Demand Model*

Following Neilson (2026), the utility for family i from school j is

$$U_{i,j} = \bar{\beta} x_j + \xi_j + \beta_i q_j - \alpha_i \text{op}_{k(i),j} + \lambda_i d_{\text{loc}(i),j} + \epsilon_{i,j}, \tag{1}$$

where q_j is school academic quality (value added), $\text{op}_{k(i),j}$ is the out-of-pocket payment for a family of type k at school j given the prevailing voucher policy, and $d_{\text{loc}(i),j}$ is distance from the family's residential location to school j. The vector x_j collects observable school characteristics over which families share common preferences $\bar{\beta}$: administration type, for-profit status, religious affiliation, and whether the school serves grades K through 12. The term ξ_j is the standard BLP-style unobserved school attribute, and $\epsilon_{i,j}$ is i.i.d. extreme value type 1. The distinction between the school's posted tuition p_j and the family's out-of-pocket payment $\text{op}_{k(i),j}$ matters for the supply side: schools set a single tuition, while out-of-pocket payments depend on how that tuition interacts with the voucher schedule; in particular, SEP-eligible families at SEP-participating schools pay zero out of pocket.[8]

Preferences are heterogeneous across six discrete family types $k \in \{1, \ldots, 6\}$, defined by mother's education and family income (Neilson, 2026). Quality preferences include a type-specific component β_k and a random coefficient $\nu_i \sim N(0, \sigma^2)$, so that $\beta_i = \beta_{k(i)} + \nu_i$, while price and distance coefficients are type-specific: $\alpha_i = \alpha_{k(i)}$ and $\lambda_i = \lambda_{k(i)}$. We group families by SEP eligibility: types $\{1, 3\}$ are poor and eligible for Chile's targeted voucher (*Subvención Escolar Preferencial*), while types $\{2, 4, 5, 6\}$ are not.

Families choose the school that maximizes $U_{i,j}$ from the set of schools in their market m. The market share of school j aggregates over types, residential locations, and the random coefficient:

$$s_j(\mathbf{q}, \mathbf{op}, \theta) = \sum_{k=1}^{K} \Pi_k^m \sum_{\text{loc}}^{L^m} w_k^{\text{loc}} \int_{\nu} s_j^{\text{loc},k}(\theta, \nu)\, d\nu, \tag{2}$$

[8] We use "location" for readability; these are the same discrete residential nodes used in Neilson (2026).

where Π_k^m is the proportion of type k families in market m, w_k^{loc} is the within-type distribution across locations, and $s_j^{\text{loc},k}(\theta, \nu)$ is the standard logit choice probability conditional on type, location, and random coefficient draw. See Neilson (2026) for the full derivation, estimation, and resulting parameter estimates $\theta = \{\bar{\beta}, \alpha, \beta, \lambda, \sigma, \xi\}$.

5.2. *Treatment Effects Parameterization*

The information treatment may affect school choice through multiple channels. Providing families with school-specific report cards can shift the weight they place on school quality, price, and distance. It can also sharpen choice behavior more broadly, increasing families' responsiveness to all school attributes. We parameterize both channels and let the data determine their relative importance.

Define the baseline (control) deterministic utility as

$$V_{ij}^C = \delta_j + \bar{\phi}_k^{q,C} q_j - \bar{\phi}_k^{op,C} \text{op}_{ij} + \bar{\phi}_k^{d,C} d_{ij} + \bar{\sigma}^C \nu_i q_j, \tag{3}$$

where $\delta_j \equiv \bar{\beta}^C x_j + \bar{\xi}_j^C$ is the school mean utility from Neilson (2026) with the quality term written separately for the purposes of the treatment parameterization, and $\bar{\phi}_k^{c,C}$ are the estimated demand coefficients from Neilson (2026) for characteristic c and type k.[9] The treated utility adds two components:

$$V_{ij}^T = \kappa_k \cdot V_{ij}^C + \psi_k^q q_j - \psi_k^{op} \text{op}_{ij} + \psi_k^d d_{ij}. \tag{4}$$

The first term rescales all baseline utility components by a common factor κ_k—including the school mean utility δ_j, the common preferences $\bar{\beta}^C x_j$, and the unobserved school attribute $\bar{\xi}_j^C$. The ψ_k^c terms add characteristic-specific treatment effects on top.

This formulation exploits an asymmetry in what the treatment targets. For school attributes about which the treatment provides no direct information—such as whether the school is public or privately administered under the voucher system—the treatment can affect choices only through the common rescaling:

$$\bar{\beta}_k^T = \kappa_k \cdot \bar{\beta}_k^C. \tag{5}$$

[9]Under the standard logit normalization, these estimates represent structural preference parameters divided by the error scale. subsection 5.3 develops this point formally; the normalization is innocuous for the baseline demand but becomes consequential when treatment may change the scale.

For attributes targeted by the information intervention—quality, out-of-pocket price, distance—the effective coefficient shifts through both the common scale and a characteristic-specific treatment effect:

$$\bar{\phi}_k^{c,T} = \kappa_k \cdot \bar{\phi}_k^{c,C} + \psi_k^c. \tag{6}$$

Equation 5 provides a key source of discipline on κ_k: for characteristics where the treatment provides no new information, the treatment effect operates only through the scale. We exploit this restriction by including moments for voucher school attendance and residualized mean utilities, both of which capture treatment effects on already-known school attributes.[10] In practice, κ_k is not pinned down by any single moment but is jointly disciplined by the full system of moments. When $\kappa_k = 1$, the model reduces to characteristic-specific treatment effects with no common scaling, nesting the standard (unscaled) specification.

Interpretation. The treatment-effect parameters ψ_k^c are summary parameters that capture the total shift in how much characteristic c influences school choice under treatment. Each ψ_k^c combines two mechanisms that are observationally equivalent in our setting. The first is an *information channel*: more precise signals raise the weight families place on the corresponding characteristic in expectation, even if underlying preferences are unchanged (Brown and Jeon, 2024). The second is a *preference channel*: by making quality differences salient and easy to compare, the report cards may cause families to place more weight on academic quality independent of any change in signal precision. A third possibility is that the intervention conveys more broadly that school quality is important for children's outcomes, potentially affecting family behavior beyond the immediate school choice—such as engagement with schooling or educational investments at other margins. This broader awareness channel could contribute to the long-run effects observed even among families who had already enrolled (Section 4), but it is not separately identified in our setting and is not modeled in the counterfactuals. The experiment identifies the combined effect of all three channels. Disentangling them would require separate data on subjective beliefs, a question we pursue in companion work (Agte et al., 2024).

For the policy counterfactuals in this paper, separately identifying these underlying channels is not necessary. The counterfactual exercises apply the same treatment observed in the

[10] The demand coefficient on voucher school status (versus public) is large and significant in Neilson (2026). Families can directly observe whether a school is public or privately administered under the voucher system, so the information treatment should not shift their assessment of this characteristic directly. Any treatment effect on the probability of attending a voucher school therefore helps discipline κ_k.

experiment to all families in the market, and simulate the resulting equilibrium. The demand shifts we estimate are identified directly from experimental variation, so the estimated total shift—however it decomposes across the information and preference channels—is the relevant sufficient statistic for these simulations.

To allow for heterogeneous treatment effects, we estimate separate ψ_k^c parameters and scale parameters κ_k for SEP-eligible families ($k \in \{1, 3\}$) and non-SEP families ($k \in \{2, 4\}$). The full treatment parameter vector is

$$\theta_2^T = \left(\psi^q_{\text{SEP}},\ \psi^{op}_{\text{SEP}},\ \psi^d_{\text{SEP}},\ \psi^q_{\text{nSEP}},\ \psi^{op}_{\text{nSEP}},\ \psi^d_{\text{nSEP}},\ \kappa_{\text{SEP}},\ \kappa_{\text{nSEP}}\right). \tag{7}$$

We incorporate the treatment into the demand model by creating *treated types* in addition to the baseline six types from Neilson (2026). The experimental sample includes families of types 1–4 (types 5–6 have too few observations in the experiment). Treated families have utility V_{ij}^T from Equation 4, while control families retain the baseline utility V_{ij}^C.

5.3. *Imperfect Information and the Scaled Logit*

The parameterization in subsection 5.2 can be interpreted through a Bayesian signal model of school characteristics. When families observe school quality, price, and distance with noise, the posterior mean for each characteristic is attenuated toward the prior by a signal-to-noise ratio $\rho_k^c \in (0, 1)$ that depends on signal precision. The effective preference weight on characteristic c becomes a product of the structural preference and the signal precision: $\phi_k^c = \beta_k^c \cdot \rho_k^c$. Crucially, as emphasized by Brown and Jeon (2024), signal noise also enters the composite error term of the choice model. Improving signal precision on specific characteristics can therefore change not only the weight on those characteristics but also the error variance of the entire choice model. We adopt a scaled logit approximation that captures this possibility while preserving the logit functional form.[11]

Dividing the expected utility by the type-specific error scale μ_k (derived in the appendix from the composite error variance) recovers standard logit choice probabilities with *effective* parameters that embed both preferences and signal precision. The baseline demand estimation in Neilson (2026) identifies only these effective parameters under the standard normalization

[11] This variance-matching approximation is standard in the heteroskedastic logit literature; see (Train, 2009), Chapters 2 and 6. It is most accurate when the EV1 taste shock dominates the signal noise, a condition supported by our empirical finding that the scale correction is modest (κ not significantly different from 1).

$\mu_k^C = 1$. This normalization is innocuous for the baseline demand but becomes consequential when treatment may change the scale. Within this framework, the common scale parameter from subsection 5.2 is

$$\kappa_k \equiv \frac{\mu_k^C}{\mu_k^T}. \tag{8}$$

When improved signal precision reduces the composite error variance, $\mu_k^T < \mu_k^C$ and $\kappa_k > 1$: the composite error shrinks, sharpening the weight on all school attributes, including those about which the treatment provides no new information. The model nests the standard logit at $\kappa_k = 1$, i.e., when treatment does not change the error variance. We do not interpret $\kappa_k > 1$ as parents literally learning about every residual school characteristic; rather, it is a parsimonious way to allow for any common sharpening of choice behavior beyond the characteristic-specific treatment shifts. Appendix A3 provides the full derivation, including the posterior, expected utility, composite error, and variance-matching formulas.

5.4. *Supply Side*

We model privately owned schools as choosing quality q_j to maximize a weighted objective that reflects both profit motives and other institutional goals. On the supply side, the relevant price variable is the school's posted tuition p_j (the sticker price). Family out-of-pocket payments op_{ij} are derived from that posted tuition after applying the voucher schedule and therefore vary by student type and school participation status. Revenue per student depends on the sticker price and the voucher schedule $v_j^k(p_j)$, where j indexes schools and k indexes student type. In the baseline demand system, posted tuition is treated as an endogenous observed school characteristic: schools that invest in unobserved quality attributes may also set higher prices, inducing a correlation between tuition and the unobserved school attribute ξ_j. The demand estimation addresses this through instrumental variables, so the recovered ξ_j already accounts for this correlation. For the supply-side exercise, we take tuition and the estimated ξ_j as given and do not model a pricing first-order condition. The reason is institutional: during the period that anchors our supply analysis, subsidized private schools operated in a regulated and segmented pricing environment rather than under standard unconstrained price setting.[12] Because our objective here is to study the medium-run quality response to the information policy, we treat observed

[12]Under the general shared-financing regime in DFL No. 2 of 1998, schools could charge positive copayments but faced caps and a reduction in the general subsidy as average fees rose. SEP then added a further layer of regulation

tuition as predetermined and focus on the quality first-order condition as the relevant adjustment margin. The per-student cost of providing quality q_j is

$$C(q_j) = \sum_l \gamma_l w_j^l + c_j \cdot q_j, \qquad c_j \equiv \gamma_q + \omega_j, \tag{9}$$

where w_j^l are observable cost shifters, γ_q is a common quality cost slope, ω_j is a school-specific unobservable cost component, and c_j is the school-specific unit cost of quality.[13] The first-order condition for optimal quality yields

$$q_j^* = \underbrace{\left[\frac{R_j^{\text{net}}}{c_j}\right]}_{\text{Competitive quality}} \underbrace{-s_j(\mathbf{q}, \mathbf{p}, \vartheta, \tau)\left[\frac{\partial s_j(\mathbf{q}, \mathbf{p}, \vartheta, \tau)}{\partial q_j}\right]^{-1}}_{\text{Quality markdown}}, \tag{10}$$

where $R_j^{\text{net}} = v_b + p_j - \sum_l \gamma_l w_j^l$ denotes net revenue per student minus observable costs (the differential revenue from SEP students enters separately through ΔR_j in the constrained FOC below). This expression decomposes optimal quality into a competitive component (quality under perfect competition) and a markdown reflecting market power. The markdown $s_j/(\partial s_j/\partial q_j)$ is inversely related to the semi-elasticity of demand with respect to quality: schools with more elastic demand face stronger competitive pressure and provide quality closer to the competitive level. Neilson (2026) derives this quality FOC and documents descriptively how the SEP reform compressed quality markdowns and raised school quality. That analysis takes advantage of observing the reform at scale, so the supply-side changes can be seen directly in the data without recovering the cost parameters. For the information policy studied here, no at-scale implementation exists, so we must recover the cost parameters $c_j = \gamma_q + \omega_j$ and solve for the counterfactual quality equilibrium. The capacity extension and revenue-composition channel described below are also new relative to that treatment.

---

by requiring participating schools to charge zero copayments to priority students (Law No. 20.248, Art. 6). See also Sanchez (2017) for related discussion of strategic school responses under Chile's voucher design.

[13]The quality index q_j recovered from the demand estimation is identified only up to differences: in the logit model, adding a constant to all q_j is absorbed by the outside-option normalization and leaves all choice probabilities unchanged. In the supply estimation, we normalize q_j so that $\min_j q_j = 0$, which ensures $(\gamma_q + \omega_j) \cdot q_j \geq 0$ for all schools. This normalization is without loss of generality for demand and guarantees non-negative costs on the supply side.

5.4.1. *Interpreting Supply-Side Parameters*

An important caveat is that the cost parameters (γ_q, ω_j) should be interpreted as a *composite* reflecting both the true resource cost of quality provision and the degree to which schools internalize objectives beyond profit. For-profit, non-profit, and public schools may all face similar technology costs for hiring better teachers, but they differ in the weight they place on revenue versus educational mission. Since we cannot separately identify the technological cost function from the weight on non-profit objectives, the school-specific cost parameter ω_j captures both margins: a school with high ω_j may face genuinely higher input costs *or* may place lower weight on profit relative to quality provision. More generally, ω_j may also absorb persistent heterogeneity correlated with the types of students a school typically serves, even though the model does not separately parameterize a marginal-cost function that varies with realized student composition. For the same reason, later accounting expenditure data are informative descriptively but do not directly observe the marginal-cost object identified by the quality first-order condition.

What the model *does* identify is the *relationship between changes in competitive incentives and the resulting quality response*. When the information treatment compresses markdowns—making demand more responsive to quality—the model predicts how much each school adjusts quality based on its estimated cost structure. This mapping from incentive changes to quality changes is well-identified regardless of whether the underlying mechanism is purely profit-driven or reflects a mix of profit and mission motives.

In practice, the supply-side model is most directly applicable to privately owned voucher schools, where quality choices reflect an optimizing response to competitive incentives and cost structures. For-profit voucher schools face the most direct revenue-quality tradeoff; non-profit voucher schools may place different weight on financial versus mission objectives, which is absorbed by the composite cost parameter but adds interpretive uncertainty. Public schools, whose quality investments are governed by municipal budgeting processes rather than school-level optimization, are the poorest fit for this framework. We therefore exclude public schools from the supply-side estimation and hold their quality fixed in the counterfactuals.

5.4.2. *Capacity Constraints and the Supply-Side FOC*

The baseline FOC in Equation 10 corresponds to the standard case in which a school can expand enrollment when higher quality raises demand. This is the same approximation that underlies the baseline demand system inherited from Neilson (2026), who argues that capacity constraints were not a first-order feature of household choice in the historical estimation

period.[14] We maintain that approximation for baseline demand estimation, but allow for the possibility that a subset of schools nonetheless faced limited room to expand enrollment on the supply side, and that such constraints become more relevant in the at-scale policy counterfactuals studied below.

Capacity-constrained schools may still have incentives to improve quality through a *composition channel*. Under Chile's Deferred Acceptance assignment mechanism, SEP-eligible students receive priority at schools that participate in the SEP program. When a school near capacity raises its quality, it need not expand total enrollment to benefit: it can instead attract a more SEP-intensive student body. Since SEP students generate higher per-student revenue ($v_b + v_{\text{SEP}}$ versus $v_b + p_j$), this compositional shift creates a financial incentive for quality improvement even when the standard enrollment channel is muted.

Formally, for a school operating at capacity $\bar{c}_j$, the first-order condition for quality becomes

$$c_j = \Delta R_j \cdot \frac{\partial s_j^{\text{SEP}} / \partial q_j}{s_j}, \tag{11}$$

where $\Delta R_j = R_j^{\text{SEP}} - R_j^{\text{nonSEP}}$ is the per-student revenue gap between SEP and non-SEP students, and $\partial s_j^{\text{SEP}} / \partial q_j$ is the derivative of the SEP-weighted share with respect to quality. The right-hand side captures the marginal revenue gain from improving composition: higher quality attracts more SEP students (who generate ΔR_j more revenue each), scaled by the fraction of the student body affected. Under the linear cost function, the marginal cost of quality is c_j (constant in q), so the FOC equates this marginal cost to the marginal composition benefit.

In practice, we classify each school-year observation as operating under one of the two regimes—unconstrained or capacity-constrained—based on observed enrollment relative to capacity, and estimate the supply parameters using both FOCs simultaneously with the school-specific cost parameter ω_j shared across regimes. Section 5.5.5 describes the classification rule and estimation details.

[14] Neilson (2026) argues in the Online Appendix that this approximation is reasonable in the baseline data because primary-level selection by voucher schools was illegal during the estimation period, only 2% of urban primary schools reached the legal class-size cap, and only 2% of surveyed parents reported preferring another school but being turned away.

5.5. *Estimation*

Our estimation proceeds in three steps: (i) recovering baseline demand parameters from Neilson (2026), (ii) estimating treatment effect parameters using the experimental variation, and (iii) recovering supply-side cost parameters.[15] We begin with an informal discussion of identification.

5.5.1. *Identification*

Baseline demand. The baseline demand parameters $\theta = \{\bar{\beta}, \alpha, \beta, \lambda, \sigma, \xi\}$ are identified by three complementary sources of variation in Neilson (2026): the BLP inversion (Berry et al., 1995) recovers school-level mean utilities ξ_j by matching predicted to observed market shares; micro-moments matching the average quality, out-of-pocket payment, and distance chosen by each family type identify heterogeneity across the six family types (Petrin, 2002); and IV moments address the endogeneity of price and quality using policy-driven revenue variation from the voucher and SEP reforms and local labor-cost shifters that proxy for the cost of hiring skilled workers. Appendix A4 lists these moments and instruments explicitly. As discussed in subsection 5.3, the baseline estimates identify the *effective* parameters that embed the normalization $\mu_k^C = 1$.

Treatment effects on informed characteristics. The ψ_k^c parameters — capturing how much treatment shifts the weight on quality, price, and distance beyond the common scale effect — are identified by the random assignment of the information treatment. Within the experimental sample, families were randomly assigned to treatment or control groups (stratified by preschool municipality). This ensures that any systematic difference in the characteristics of schools chosen by treated versus control families is caused by the treatment, not by selection. The moments match these experimental treatment effects, separately for each characteristic and type group.

Scale parameter. The scale parameters κ_g are jointly disciplined by the full system of 12 moments. Treatment effects on voucher school attendance help pin down κ_g because voucher status is an already-known characteristic: any model fit to those moments must come through the common rescaling channel (Equation 5) rather than through a direct information-treatment shift. Similarly, moments based on residualized mean utilities $\tilde{\delta}_j \equiv \delta_j - \bar{\beta}^q q_j$ capture treatment effects on the composite of already-known school attributes embedded in the BLP mean utility.

[15] See Appendix A4 for additional estimation details and standard error formulas.

The treatment effects on informed characteristics (q, op, d) also contribute, since the model attributes part of the observed shifts to the common scale effect and part to the characteristic-specific ψ_g^c parameters. No single moment identifies κ_g in isolation; it is the full system that disciplines the 8-parameter vector.

Supply. Supply parameters are identified from the first-order conditions for quality (Equation 10 and Equation 11). We rearrange these FOCs to express the "competitive quality" residual—the gap between observed quality and the quality markdown—as a linear function of cost shifters, school fixed effects, and market-year fixed effects. The SEP indicator serves as the key cost shifter w_j^l, capturing how participation in the targeted voucher program affects the cost or incentive structure. Cross-sectional variation in the SEP indicator and the time-series variation from the policy's staggered rollout provide the instruments. The school-specific cost ω_j is concentrated out analytically, reducing the dimensionality of the GMM problem. Identification of γ_q relies on the cross-school and cross-time variation in markdowns, which are themselves functions of the demand-side parameters from Steps 1 and 2.

5.5.2. *Step 1: Baseline Demand Parameters*

The baseline demand parameters are taken from Neilson (2026), who estimates the first-step demand system using administrative first-grade choice data from 2006, 2007, 2011, and 2012. We refer the reader to that paper for a detailed description of the estimation strategy and the resulting parameter estimates.

5.5.3. *Estimation Sample*

The treatment effect parameters are estimated using a subsample of the experimental families that can be matched to the market-level demand model. We begin with the full experimental sample and apply a series of restrictions, each motivated by the requirements of the structural estimation. First, we restrict attention to families that had *not yet enrolled* their children at the time of the Pre-K visit, since the treatment effects on school choice are concentrated entirely among these families (section 4). Second, we restrict to the six urban markets that overlap between our experiment and the market structure in Neilson (2026), which is necessary because the baseline demand parameters—mean utilities δ_j, quality measures, and spatial choice sets—are only available within Neilson's market definitions. Third, we apply a geographic consistency filter: we keep families if they remain in the same municipality as at baseline or if the chosen

school is within 4 km of the family's home, excluding only observations that are both in a different municipality and implausibly far from the chosen school.

We then drop families with missing information on the school they ultimately chose (value-added, out-of-pocket price, or distance). Finally, family types are assigned using six categories defined by mother's education and SEP eligibility; because the experiment contains few families of types 5–6 (technical and college-educated), we collapse these into types 3–4 using actual JUNAEB SEP eligibility records to ensure correct classification.

Table III reports the treatment effects on chosen school characteristics in the resulting estimation sample, separately for SEP-eligible and non-SEP families. These are the experimental moments that the structural model must match. The treatment significantly increases the value-added of schools chosen by SEP families (0.220, $p < 0.05$); the point estimate on driving distance is also positive (0.134 km) though not statistically significant. These treatment effects form the empirical targets for the GMM estimation described below.

TABLE III

TREATMENT EFFECTS IN THE MODEL SAMPLE (EXPERIMENT MOMENTS)

	Quality	Sticker Price	OOP Price	Distance	Voucher	N
Pooled	0.195***	0.039**	0.027	0.118	0.114***	815
	(0.073)	(0.017)	(0.017)	(0.078)	(0.042)	
SEP	0.220**	0.025	0.006	0.134	0.106**	505
	(0.087)	(0.017)	(0.016)	(0.092)	(0.043)	
Non-SEP	0.114	0.048	0.048	0.184	0.096	310
	(0.104)	(0.036)	(0.036)	(0.119)	(0.067)	
Preschool center FE	Yes	Yes	Yes	Yes	Yes	

Notes: Experimental moments used in the demand estimation. Regressions use the model-overlap sample described in Section 5.5.3, preschool-municipality fixed effects, and standard errors clustered at the preschool-center level; exact controls are summarized in Appendix A2. Quality is normalized value added Q_2. Distance is Google Maps driving distance, restricted to observations within 4 km. Voucher indicates a private voucher rather than a public school. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.5.4. *Step 2: Treatment Effect Parameters*

Conditional on the Step 1 estimates, we estimate the treatment parameter vector θ_2^T defined in Equation 7 using GMM. The identification strategy exploits the random assignment of the information treatment within the experimental sample.

We define *RCT moments* that match the treatment effects on school characteristics observed in the experimental data. For each characteristic $c \in \{q, p^{\text{sticker}}, \text{op}, d\}$ and each type group $g \in \{\text{SEP}, \text{nonSEP}\}$, the moment condition requires the model-simulated treatment effect to match the experimental estimate:

$$G_{c,g}(\theta_2^T) = \hat{\beta}_{c,g}^{\text{RCT}} - \hat{\beta}_{c,g}^{\text{sim}}(\theta_2^T), \tag{12}$$

where $\hat{\beta}_{c,g}^{\text{RCT}}$ is the coefficient on the treatment indicator from a regression of characteristic c on treatment status (with stratification fixed effects) in the experimental data, and $\hat{\beta}_{c,g}^{\text{sim}}(\theta_2^T)$ is the analogous coefficient computed from model-simulated school choices. The simulated choices are generated using the treated and control probabilities implied by the model, with the random coefficient integrated via Gauss-Hermite quadrature. To help discipline the scale parameters κ_g, we add moments on voucher-school attendance and residualized mean utilities $\tilde{\delta}_j \equiv \delta_j - \bar{\beta}^q q_j$, both of which capture treatment effects on already-known school attributes. Because these characteristics were not targeted by the information intervention, any model fit to these moments must come through the common rescaling channel (Equation 5) rather than through direct information shifts. Together, the full system of moments jointly disciplines all eight parameters (seven freely estimated in the benchmark specification, with $\psi_{\text{SEP}}^{\text{op}}$ restricted to zero as discussed in the results section).

We estimate using two-step efficient GMM: the first step uses the identity weighting matrix, and the second step uses $W = \hat{V}_{\text{CR}}^{-1}$, where $\hat{V}_{\text{CR}}$ is the cluster-robust covariance matrix of the data moments (clustering at the preschool center level). Standard errors are computed using the standard GMM sandwich formula.

5.5.5. *Step 3: Supply Parameters*

Given the demand parameters from Steps 1 and 2, we estimate the supply-side cost parameters using GMM on the quality first-order conditions. We classify a school-year observation as capacity-constrained when its enrollment-to-capacity ratio exceeds 0.95 and the school receives differential SEP revenue. This classification uses observed enrollment and imputed capacity and should therefore be understood as a transparent empirical rule rather than an exact measure of when capacity begins to bind. In our preferred specification, 2,414 of 22,882 school-year observations (10.5%) are classified as constrained, so the two-regime extension modifies the supply problem for a meaningful but limited subset of the sample.

For each school-year observation, we compute the quality markdown $\text{MD}_{jt} = s_{jt}/(\partial s_{jt}/\partial q_j)$ from the estimated demand system and define the "competitive quality" residual:

$$\text{CompQ}_{jt} \equiv q_{jt} + \text{MD}_{jt} = \frac{R_{jt}^{\text{net}}}{c_j} + \gamma_{mt}^{\text{type}}, \tag{13}$$

where $c_j = \gamma_q + \omega_j$ is the school-specific unit cost of quality and $\gamma_{mt}^{\text{type}}$ are market-year-type fixed effects that absorb common cost shocks separately for voucher and private non-voucher schools. This is the left-hand-side object implied by rearranging the unconstrained quality FOC; it is not a direct measure of "quality investment" or of mean utility. For concentration, however, we work with the inverse transformation $A_{jt} \equiv R_{jt}^{\text{net}}/\text{CompQ}_{jt}$, where the cost terms enter linearly, as shown in Appendix A4. For capacity-constrained observations, we use the analogous residual from Equation 11, which replaces the enrollment-based markdown with the composition-based incentive.

We concentrate out c_j and $\gamma_{mt}^{\text{type}}$ analytically. The cost function is parameterized as $c_{jt} = \gamma_{\text{SEP}} \cdot \text{SEP}_j + \delta_{\text{type}(j)} \cdot t + c_j$, where γ_{SEP} captures the cost differential associated with SEP program participation and δ_{type} are type-specific time trends that allow secular changes in cost structures to differ across school types. The school-specific cost c_j is the time-average residual after removing these components. The estimation sample consists of voucher and private non-voucher schools observed over 10 years (2006–2015) across 53 markets (22,882 school-year observations, 2,389 schools). Public schools are excluded for the reasons discussed in Section 5.4.1.

The moment conditions interact the structural residual $\Omega_{jt} \equiv \text{CompQ}_{jt} - \hat{\text{CompQ}}_{jt}$ with instruments. We use demand-side exclusion restrictions interacted with school type, plus a Nearby Teacher Quality Index — the inverse-distance-weighted average of public school teacher PSU scores within each market — which captures cost-shifting pressure from the teacher labor market without directly affecting school demand. Table O-11 reports the corresponding first-stage relationship using CompQ_{jt} as the dependent variable. We estimate using two-step optimal GMM. We view this two-regime supply specification as a disciplined extension of the baseline model rather than a fully unified model of demand, assignment, and capacity: it uses the same latent-demand primitives as the rest of the paper, but lets the marginal incentive to raise quality depend on whether a school appears to have slack enrollment capacity.

TABLE IV

DEMAND MODEL ESTIMATES

	Panel A: Baseline Demand (Neilson 2026)				*Panel B: Treatment Effects (this paper)*		
	Quality (β^q)	Price (α)	Distance (λ)		Quality (ψ^q)	Price (ψ^{op})	Distance (ψ^d)
<HS	1.40 (0.03)	−2.70 (1.10)	−1.29 (0.05)	SEP	0.500 (0.230)	—	0.427 (0.204)
HS	0.56 (0.10)	−0.56 (0.38)	−1.09 (0.04)	nonSEP	0.399 (0.289)	−0.148 (0.067)	0.061 (0.269)
Tech/2-Year	0.86 (0.16)	−0.25 (0.38)	−1.04 (0.04)				
College+	1.13 (0.22)	−0.00 (0.38)	−0.96 (0.04)		*Panel C: Scale Parameters*		
SEP	−0.33 (0.08)	−1.47 (0.31)	−0.06 (0.02)	κ_{SEP}	1.410 (0.291)		
				κ_{nonSEP}	1.258 (0.492)		
Rand. coeff. (σ_q)	0.85 (0.31)						

Notes: Panel A reports baseline demand parameters from Neilson (2026); quality coefficients are deviations from $\bar{\beta}^q$, with less-than-high-school mothers as the omitted category. Panel B reports treatment-effect parameters estimated here using the sample in subsubsection 5.5.3. Panel C reports scale parameters; $t(\kappa{=}1)$ tests the null of no scale change. Standard errors in parentheses are clustered at the preschool-center level. $\psi^{\text{op}}_{\text{SEP}}$ is set to zero because SEP-eligible families face zero out-of-pocket prices at SEP schools, making this parameter unidentifiable (see Online Appendix for the unrestricted specification).

5.6. *Parameter Estimates*

Table IV presents the demand-side parameter estimates. Panel A reports the baseline demand parameters from Neilson (2026); Panel B reports the treatment effect parameters estimated in this paper. The baseline model is defined over six family types based on mother's education and SEP eligibility, while the experimental treatment heterogeneity is summarized more parsimoniously by SEP status because the intervention sample is concentrated in the lower-education groups targeted by the policy. We refer the reader to Neilson (2026) for a thorough discussion of the baseline demand estimates and focus here on the new treatment parameters.

The main demand-side result is that the information treatment increases the weight families place on school quality. This effect is clearest for SEP families, where the estimated shift is economically meaningful relative to the baseline quality coefficient, but the pattern is similar for non-SEP families as well. In other words, the treatment does not simply move families mechanically toward particular schools; it makes school quality matter more in their choices.

The other treatment parameters point in the same direction. For SEP families, treatment lowers the effective penalty of distance, consistent with greater willingness to look beyond the nearest options when higher-quality schools are available. For non-SEP families, the treatment makes out-of-pocket price matter somewhat more. Taken together, these estimates line up with the descriptive experimental evidence from Section 4: treated families place more weight on academic quality and are more willing to trade off convenience, and in some cases price, when

choosing schools. Because SEP-eligible families face zero out-of-pocket prices at SEP schools by construction, the corresponding price channel is not separately identified for that group.

The scale estimates provide only modest evidence that the intervention changed the overall sharpness of choice behavior. Both κ estimates are above one, which is the direction implied by the imperfect-information interpretation, but only the SEP estimate is marginally distinguishable from the benchmark of one. The broader message is therefore not that the treatment primarily worked by dramatically reducing global choice noise. Rather, the data suggest that the main action comes through the characteristic-specific shifts, especially the increased weight on quality. Allowing for the scaled-logit possibility emphasized by Brown and Jeon (2024) is still useful conceptually, but in this application it does not materially overturn the main demand-side interpretation.

5.7. *Supply Model Estimates*

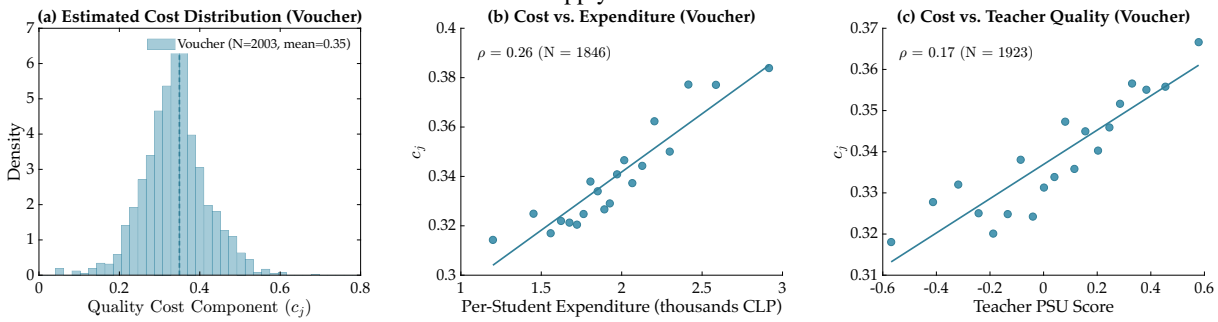
Figure 2 presents the supply-side cost estimates from the two-regime GMM procedure described in subsubsection 5.5.5, estimated on voucher and private non-voucher schools over 10 years (2006–2015) across 53 markets.[16] The estimation concentrates out school fixed effects c_j and market-year-type fixed effects $\gamma_{mt}^{\text{type}}$ analytically, with type-specific time trends $\delta_{\text{type}} \cdot t$ allowing secular changes in cost structures across school types. Identification comes from demand-side exclusion restrictions interacted with school type, plus a Nearby Teacher Quality Index that captures cost-shifting pressure from the teacher labor market (Table O-11).

The supply estimates indicate that schools participating in the targeted voucher program face meaningfully different cost structures, and the two-regime specification is empirically relevant rather than merely conceptual: a nontrivial share of school-year observations is classified as constrained and therefore governed by the composition-channel first-order condition rather than the standard enrollment-based one. The time trends by school type are comparatively modest, so the main variation of interest comes from the cross-school differences in competitive incentives and the associated cost structure.

---

[16]The demand parameters are estimated on 2006, 2007, 2011, and 2012 as described in subsubsection 5.5.2. The extended supply panel uses the same demand estimates and recovers mean utilities for the additional years (2008–2010, 2013–2015) via BLP contraction mapping. Public schools are excluded from the main specification because their quality investment decisions are governed by municipal budgeting processes rather than the school-level optimality conditions in the model.

FIGURE 2.—Supply-Side Cost Estimates



Notes: Panel A: distribution of estimated school-specific quality cost $\hat{c}_j$ for voucher schools; dashed line marks the mean. Panels B and C: bin-averaged $\hat{c}_j$ against per-student expenditure and average teacher PSU scores for voucher schools. 22,882 school-year observations (2,389 schools), 53 markets, 2006–2015. Full parameter estimates in Table O-10.

Panel A of Figure 2 shows the distribution of estimated quality costs for voucher schools. The distribution is concentrated at low values, with a mean of 0.35 and a tight dispersion. The most useful interpretation is not that the model recovers a pure technological marginal cost curve, but that it recovers a composite object summarizing how costly it is for different schools to raise quality, given both resource costs and institutional objectives. Panels B and C provide a simple validation of that interpretation: schools that the model treats as having higher quality costs also tend to spend more per student and employ teachers with stronger observed credentials.[17] That is exactly the pattern one would expect if the estimated cost term is capturing a meaningful constraint on quality improvement rather than merely residual noise.

In the next section, we use these demand and supply estimates to study how the policy would operate at scale under alternative assumptions about assignment, capacity, and equilibrium responses.

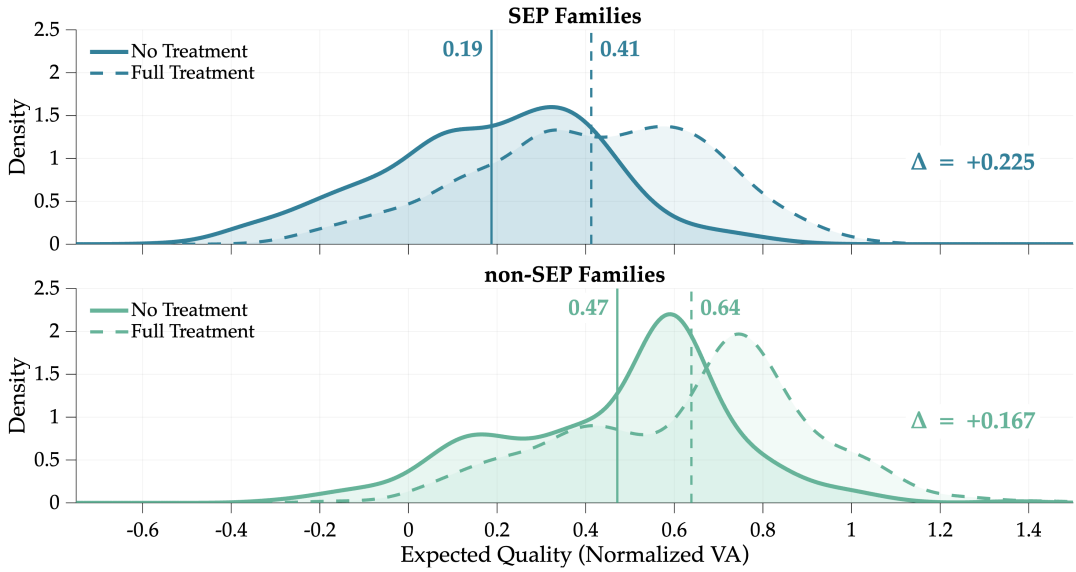
6. EQUILIBRIUM EFFECTS

6.1. *Counterfactual Policy Simulations*

We now use the estimated model to ask what would happen if the information intervention were implemented at scale rather than in a small experiment. We build up the answer in layers, isolating the contribution of each equilibrium force. The first benchmark holds school characteristics fixed and asks how demand shifts alone would change the distribution of school quality. The second introduces capacity constraints through centralized assignment. The third allows

[17] Per-student expenditure comes from MINEDUC's accounting data (*Gasto Rendido*) and includes both personnel and operational costs.

FIGURE 3.—Counterfactual: Quality Distribution under Universal Information (Experiment Sample)



Notes: Distribution of expected school quality (normalized VA) for experiment-sample families. Counterfactual applies the information treatment to all families ($\tau = 1$), holding school quality and prices fixed. Top panel: SEP families. Bottom panel: non-SEP families. Solid lines: baseline (no treatment). Dashed lines: full treatment. Vertical lines mark group means.

schools to adjust quality in response to the changed competitive environment. This layered approach makes clear which forces matter and in which direction.

Formally, let $\tau_{nk} \in [0, 1]$ denote the fraction of type-k families at residential location n who receive the treatment, following the spatial structure of Neilson (2026). Each family's school choice probabilities are a mixture of the treated and control predictions from the demand model. Aggregating these mixed shares across locations yields market-level demand by type; the average quality attended by each type is the corresponding share-weighted mean over schools.

Model Fit and At-Scale Predictions Before exploring the layered counterfactuals, we verify that the structural model reproduces the experimental evidence and then examine how the implied treatment effect changes when extrapolated to the broader population. Table V reports three objects for each family group. Column 1 is the experimental RCT average treatment effect on school quality from Table III. Column 2 is the model-predicted treatment effect computed over the same experiment-sample markets—an in-sample fit check confirming that the structural parameters reproduce the targeted moments when aggregated through the counterfactual machinery. Column 3 is the model-predicted effect computed over all 53 urban markets in the model population. The difference between columns 2 and 3 is informative: for SEP families, the population effect (0.185 SD) is somewhat smaller than the experiment-sample prediction

TABLE V

SUMMARY OF DEMAND-SIDE POLICY EFFECTS

	Panel A: Treatment Effects on Quality			*Panel B: Quality Gap*	
	Experiment	Model (demand only)		No treatment	−0.214
	RCT ATE	Exp. sample	All mkts	Full treatment	−0.173
SEP families	0.220	0.221	0.185	Narrowing	0.041 (19%)
Non-SEP families	0.114	0.151	0.144		

Notes: Panel A reports *treatment effects*—the difference in mean school quality between full treatment ($\tau = 1$) and no treatment ($\tau = 0$), holding school characteristics fixed. Column 1 is the experimental RCT ATE from Table III. Column 2 is the model prediction in the experiment-sample markets. Column 3 is the model prediction across all 53 urban markets; this is the demand-only benchmark carried forward in Table VI, Column (2). Panel B reports the SEP–non-SEP quality gap under no treatment and full treatment across all markets. Quality is measured in standard deviations of normalized value added.

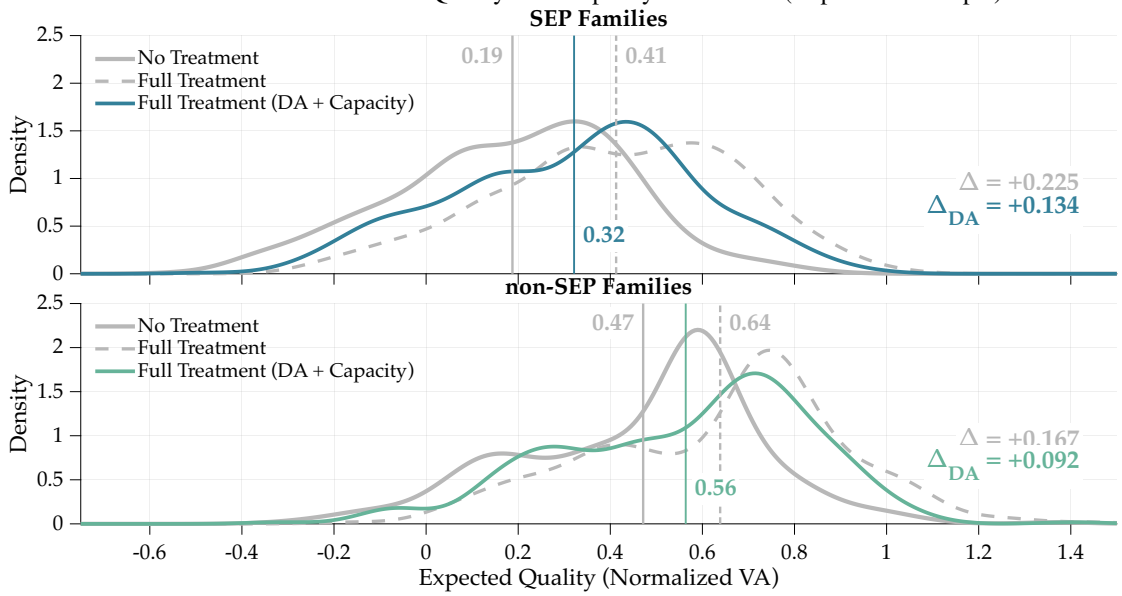
(0.221 SD). This is consistent with the experiment operating in markets where low-income families face larger quality gaps and therefore have more room to gain from information. The population extrapolation is the relevant benchmark for the policy simulations that follow.

At-Scale Demand Responses As a first benchmark, we ask what the treatment would imply if schools had not yet adjusted—the natural partial-equilibrium starting point. We apply the treatment to all families ($\tau = 1$) while holding school quality and prices fixed, and compute the implied distribution of school quality for SEP and non-SEP families (Figure 3). The demand shift produces a substantial rightward movement in quality for SEP families, reflecting the large quality weight increase estimated in Table IV.

However, these demand shifts imply that the most popular schools would have to expand enrollment substantially—a requirement that may not be feasible in the short run. This does not contradict the baseline demand approximation inherited from Neilson (2026), where capacity constraints are not a first-order concern in the historical data. The scale-up exercise asks a different question: once the information policy is applied broadly, demand can become sufficiently concentrated at popular schools that enrollment limits matter for realized assignments even if they were second-order in the estimation sample.

Capacity-Constrained Counterfactual The demand-only benchmark assumes every family can attend their most-preferred school. In practice, popular schools face enrollment limits. To assess how congestion attenuates treatment effects, we simulate a student-proposing Deferred Acceptance (DA) algorithm in each market, where families submit rank-ordered preference

FIGURE 4.—Counterfactual: Quality with Capacity Constraints (Experiment Sample)



Notes: Gray solid: baseline. Gray dashed: full treatment, free choice. Colored solid: full treatment under DA with capacity constraints. Treatment applied to types 1–4; types 5–6 untreated.

lists derived from their logit choice probabilities and schools rank applicants by random lottery priority, as in Chile's centralized system.[18]

School capacities are imputed from administrative records. For public and voucher schools we use reported enrollment slots (*cupos totales*) from Chile's 2020 school-level enrollment data (SAE), adjusted to the estimation period using historical classroom counts (see Online Appendix O-13.2 for details). For unmatched and private schools we use alternative administrative capacity measures and market-level medians as fallbacks. The resulting enrollment-to-capacity ratio averages 0.73, indicating meaningful slack in most markets. The DA exercise should therefore be interpreted as a short-run congestion benchmark, not as evidence that household choices in the baseline estimation sample were broadly rationed.

Figure 4 plots the quality distributions under three scenarios: baseline, demand-only (free choice), and demand under DA with capacity constraints. Table VI decomposes the equilibrium effects by reporting *mean quality levels* across the population under each successive layer—the treatment effects in Panel B are the implied differences relative to the no-treatment baseline. Column (2) corresponds to the demand-only population benchmark from Table V, Column 3; the remaining columns progressively add capacity constraints and supply responses. Conges-

[18]To generate rank-ordered preference lists consistent with the multinomial logit model, we draw $\tilde{U}_{ij} = V_{ij} + \varepsilon_{ij}$ for each family-school pair, with ε_{ij} i.i.d. Gumbel, and then rank schools by $\tilde{U}_{ij}$.

TABLE VI

COUNTERFACTUAL SIMULATIONS: MEAN QUALITY AND TREATMENT EFFECTS (POPULATION)

			Equilibrium (DA + Capacity)		
	(1)	(2)	(3)	(4)	(5)
	Baseline	Demand Only	No Supply Response	Supply Response	Supply Resp. (+10% MC)
Panel A: Mean Expected Quality (SD)					
All (types 1–6)	0.558	0.680	0.619	0.723	0.615
SEP (types 1, 3)	0.298	0.483	0.407	0.526	0.444
Non-SEP (types 2, 4)	0.512	0.656	0.588	0.702	0.599
Panel B: Treatment Effects (relative to Baseline)					
All (types 1–6)		0.122	0.062	0.166	0.057
SEP (types 1, 3)		0.186	0.109	0.229	0.147
Non-SEP (types 2, 4)		0.144	0.076	0.191	0.087

Notes: Panel A reports population-weighted mean expected quality across all markets; Panel B reports the implied treatment effects (each column minus Column 1). Column (2) is the demand-only benchmark from Table V, Column 3, reported here as a level to facilitate comparison with the equilibrium columns. Columns (3)–(5) impose deferred acceptance with administrative capacity constraints. "No Supply Response": schools do not adjust quality. "Supply Response": schools adjust quality via the estimated FOC; public schools held fixed. Column (5) scales all schools' unit cost of quality c_j by 1.10 before solving for the equilibrium, capturing potential upward pressure on input costs when many schools improve quality simultaneously (Tincani, 2020). Capacities from administrative data (Online Appendix O-13.2).

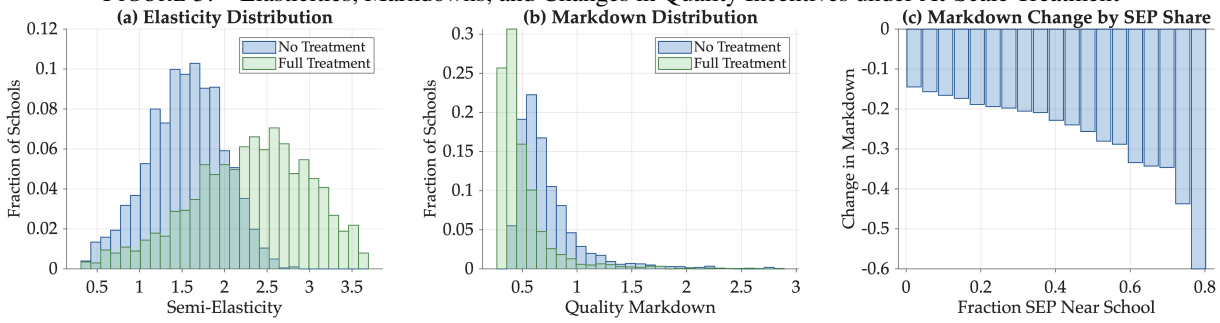
tion matters: roughly half of the demand-only treatment effect survives capacity constraints (Column 3). Untreated higher-education families (types 5–6) experience a small negative displacement as treated families crowd into higher-quality schools—a spillover that the supply response in the next section can reverse.

6.2. *Supply Responses*

The capacity-constrained simulations show that congestion attenuates roughly half of the demand-side gains when the policy is scaled up. The central question is whether supply-side responses—schools adjusting quality in response to stronger competitive pressure—can offset this attenuation. We interpret the exercise as a medium-run approximation: schools adjust quality given the shifted demand system and observed capacity, but we abstract from entry, exit, and capacity investment.

Before computing the equilibrium, it is useful to see how the information treatment changes competitive incentives on impact. Figure 5 presents three panels. Panel (a) shows the distribution of the quality semi-elasticity of demand under baseline and full treatment: the median increases

FIGURE 5.—Elasticities, Markdowns, and Changes in Quality Incentives under At-Scale Treatment



Notes: Demand-side only (no supply adjustment), all families treated ($\tau = 1$). Panel (a): distribution of the quality semi-elasticity of demand $(\partial s_j/\partial q_j)/s_j$ for voucher schools under baseline and full treatment. Panel (b): distribution of quality markdowns $s_j/(\partial s_j/\partial q_j)$ under baseline and full treatment. Panel (c): median change in markdown by the fraction of SEP-eligible students near the school. 53 urban markets.

from 1.55 to 2.37, a 53% increase. Panel (b) shows the corresponding quality markdowns, which compress from a median of 0.65 to 0.42—a 35% reduction in market power over quality. Panel (c) shows that this compression is largest for schools in neighborhoods with the highest concentration of SEP-eligible students, consistent with the treatment having the greatest impact on the families that were previously least informed. The policy therefore does not just shift demand; it also changes the competitive environment in which schools operate, especially in the neighborhoods where baseline inequality is greatest.

These changes in competitive incentives motivate the equilibrium exercise. We solve for the quality vector that satisfies each school's first-order condition under the shifted demand system. For unconstrained schools, the equilibrium update simplifies to $Q^{\text{eq}} = Q^{\text{data}} + \text{MD}^{\text{base}} - \text{MD}^{\text{treat}}(Q^{\text{eq}})$, so the quality response is governed by how markdowns change. For capacity-constrained schools, the composition-channel FOC (Equation 11) determines how much quality adjustment the estimated cost structure and the SEP revenue gap can support. Both regimes use the estimated supply parameters: the cost estimates anchor the constrained-school response directly and determine the sensitivity of the equilibrium to input cost pressure through the robustness exercise described below.

We iterate the two-regime fixed-point equation—updating demand, recomputing markdowns and composition-channel derivatives, and adjusting quality with damping—until convergence. Capacity status is classified dynamically at each iteration based on implied enrollment. Schools facing larger reductions in markdowns increase quality more. Public schools are held fixed ($\Delta q_j = 0$), making the exercise conservative on the public-school margin.

Figure 6 shows the result. The supply response substantially amplifies the treatment effect, pushing the equilibrium distribution well beyond the demand-only benchmark. For SEP families in the experiment sample, the combined demand-plus-supply treatment effect under DA is $\Delta_{DA+S} = +0.257$ SD, compared to $\Delta_{DA} = +0.134$ with demand only under DA and $\Delta = +0.225$ under free choice. The supply amplification more than compensates for the attenuation from capacity constraints: the equilibrium effect exceeds the free-choice demand-only benchmark, meaning that a partial-equilibrium analysis would understate the policy's impact. Non-SEP families experience similar patterns ($\Delta_{DA+S} = +0.208$ versus $\Delta_{DA} = +0.092$). At the population level (Table VI), the equilibrium treatment effect exceeds the demand-only benchmark by a factor of 1.23–1.36 across family types. This pattern is also robust to upward pressure on input costs. Even when the unit cost of quality is scaled up by 10%, the equilibrium treatment effect for SEP families remains positive at $+0.147$ SD; Figure O-9 in the appendix shows the full sensitivity analysis for cost increases up to 30%.

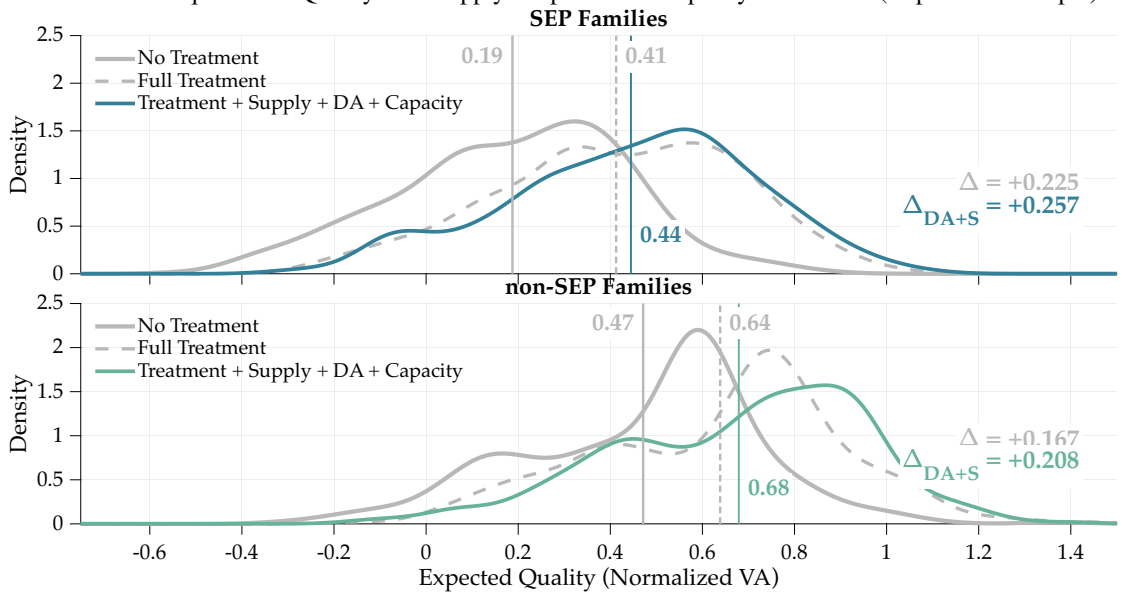
The central lesson is that partial-equilibrium demand calculations can mislead in opposite directions depending on which equilibrium margins are omitted. A naive extrapolation that ignores congestion would overstate the short-run gains. But an analysis that accounts for congestion without allowing supply responses would understate the medium-run gains. The equilibrium analysis reveals that the net effect of scaling up is positive and larger than the demand-only benchmark, because the competitive channel—schools facing stronger incentives to improve quality—dominates the congestion channel.

6.3. *Robustness and Discussion*

The equilibrium exercises rest on several modeling choices and institutional assumptions. We discuss the most important ones here; Table O-13 in the appendix reports additional robustness exercises—including targeted treatment of only SEP or non-SEP families, allowing public schools to respond, and partial take-up—which confirm that the main conclusions are stable and that supply-side quality improvements generate positive spillovers for untreated families.

Over a longer horizon, entry, exit, or capacity investment could in principle respond to a sustained demand shift. In practice, prior work in the Chilean context finds that these margins have not been large drivers of market-level change, given the existing market structure and apparent excess capacity (Neilson, 2026). Similar to the analysis in Wollmann (2018), our counterfactual focuses on the adjustment of product characteristics by incumbent schools. If anything, binding capacity constraints and regulatory barriers to entry dampen competitive pressure, making

FIGURE 6.—Equilibrium Quality with Supply Response and Capacity Constraints (Experiment Sample)



Notes: Distribution of expected school quality (normalized VA) for experiment-sample families under three scenarios. Gray solid: baseline (no treatment). Gray dashed: full treatment, demand only (no supply response). Colored solid: full treatment under deferred acceptance with administrative capacity constraints and supply-side quality response. Top panel: SEP families. Bottom panel: non-SEP families. Public schools held fixed ($\Delta q = 0$).

our supply-response estimates conservative relative to a longer-run equilibrium in which these margins are also active.

The counterfactual also leans on two institutional features of the Chilean environment. First, the assignment mechanism is centralized deferred acceptance, which limits schools' ability to select students strategically when demand shifts. Second, copayments were already heavily regulated during our estimation period and were subsequently frozen under the 2015 Inclusion Law, so schools cannot absorb increased demand by raising prices.[19] Together, these features justify focusing on quality as the medium-run adjustment margin and abstracting from strategic pricing and admissions responses.

7. DESIGNING A LARGE-SCALE RCT

The preceding sections established that equilibrium forces—congestion and supply-side quality adjustment—are quantitatively important for evaluating the information policy at scale, and that an individual-level RCT cannot capture them: if treating one family shifts demand at a

[19]Even before the 2015 law, the shared-financing rules (DFL No. 2 of 1998) capped copayments and reduced the general subsidy as fees rose; SEP required participating schools to charge eligible students zero. See Sanchez (2017) for related discussion.

school, nearby control families are affected too. Cluster-randomized trials—which assign entire geographic areas to treatment or control—can recover such effects by comparing outcomes across differentially saturated clusters (e.g., Miguel and Kremer, 2004, Muralidharan and Sundararaman, 2015). In school choice, the natural cluster would be the market, but with only 53 urban markets available in this setting, market-level randomization may lack statistical power. The model, however, reveals that competitive pressure is highly location-specific: a school's demand depends primarily on nearby families, and the supply response depends on local markdown compression. This locality means that spillovers decay with distance, making it possible to create geographic clusters *within* a single market—small enough to yield many independent units for estimation, yet large enough to contain the relevant competitive interactions. The remaining design problem is contamination: equilibrium responses propagate through competition among geographically connected schools, so schools near cluster boundaries are not independent. At the same time, having more clusters improves statistical power.

We use the estimated model to evaluate candidate cluster-randomized designs and study their statistical properties before the experiment is fielded. We focus on Santiago, the market where the original experiment was conducted (Figure 7, Panel a), and study how cluster size, treated area, and buffer zones shape the bias and precision of the design.

Design. We tile Santiago with a regular hexagonal grid of circumradius r km (Figure 7).[20] Each eligible hexagon is assigned to treatment or control via stratified randomization, balanced 50/50 within strata defined by distance from the market centroid and angular quadrant.

Within each hexagon, two concentric zones serve different roles. The *core* (distance $< I$ km from center) is the unit of observation: only schools in the core zone enter the regression. The surrounding *treatment ring* (distance I to r) extends the treated area so that core schools face high demand saturation from treated neighbors, but ring schools are excluded from the regression to avoid noisy near-boundary observations. In treated hexagons, one quarter of families receive the information intervention ($\tau = 0.25$); in control hexagons no one does.[21]

Estimand and estimator. The estimand is $\beta^* = \bar{Y}(\tau{=}\mathbf{1}) - \bar{Y}(\tau{=}\mathbf{0})$, the average treatment effect on school quality at the specified take-up rate. The outcome is the school-level quality change induced by the supply-side response, ΔQ_j, for voucher and private schools in the core

[20] Hexagons approximate circles—minimizing the boundary-to-area ratio—while tiling the plane without gaps, making them a natural choice for spatial cluster designs.

[21] The appendix examines sensitivity to higher take-up rates.

zone—the schools whose quality adjusts via the supply-side FOC.[22] The estimator is

$$\Delta Q_j = \alpha + \beta\, T_{h(j)} + \gamma\, Q_{j,\text{baseline}} + \epsilon_j, \tag{14}$$

where ΔQ_j is the quality change at school j in the core of hexagon h, $T_{h(j)}$ is the treatment indicator for that hexagon, and $Q_{j,\text{baseline}}$ is pre-treatment quality. Standard errors are clustered at the hexagon level.[23]

Bias–power trade-off. The fundamental trade-off is between contamination bias and statistical power. Smaller clusters share more schools across treatment-control boundaries (increasing the potential for contamination and downward bias), while larger clusters leave fewer independent units for estimation (decreasing power). Because treatment and control clusters share schools near their boundaries, the estimator is attenuated. A useful heuristic is the exposure contrast $\rho = \bar{\lambda}^{\mathcal{T}} - \bar{\lambda}^{\mathcal{C}}$, where λ_j is the treated share of school j's demand: in a simple difference-in-means estimator, bias would be approximately $-(1-\rho)\beta^*$. In practice, the three-zone design reduces contamination well below this bound by excluding near-boundary schools from the regression. We optimize directly over the simulation-based RMSE, computed across stratified random assignments rather than inferred from the analytic approximation.[24]

Simulation. We evaluate a grid of cluster designs, varying the hex radius from 0.75 to 4 kilometers and the core radius from 0.5 to 2 kilometers. For each configuration, we run multiple simulation replicates—each drawing structural parameters from their asymptotic distribution, a random hex grid origin (to average over grid placement), and a stratified hex assignment—and estimate (14) on Santiago clusters. The outcome is the equilibrium quality response: we solve the supply-side FOC under partial treatment and measure how much each school's quality changes relative to the no-treatment baseline ($\beta^* \approx 0.05$ s.d. at $\tau = 0.25$).[25]

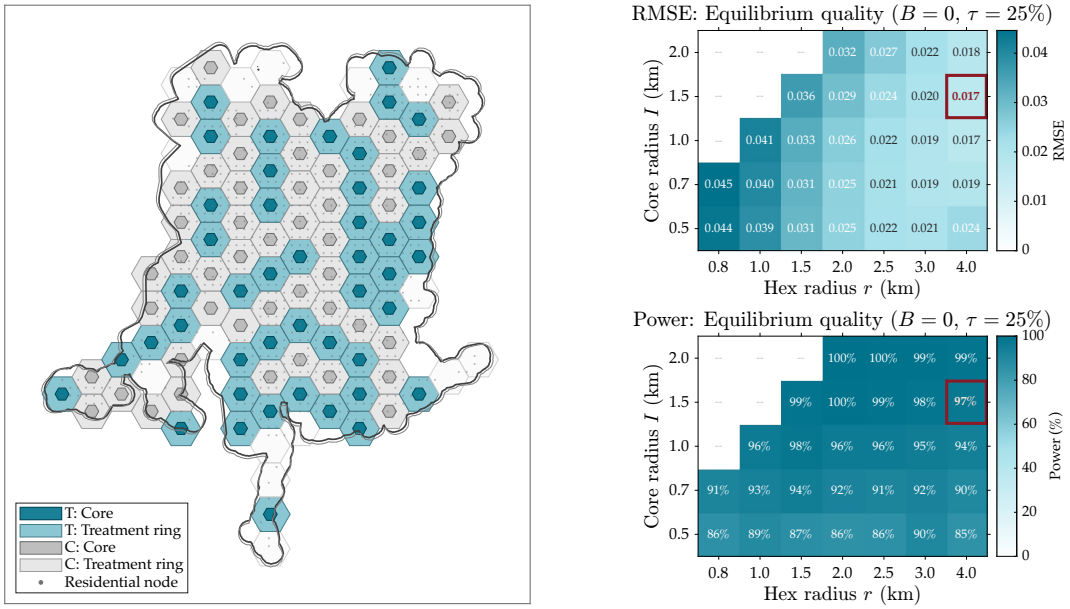
[22] Public schools are excluded from the regression because they do not respond to competitive incentives through the model's supply channel.

[23] Because the design is restricted to Santiago, the regression does not include market fixed effects. In a multi-market implementation, market fixed effects would absorb between-market variation in school quality.

[24] Each simulation replicate draws treatment parameters from their asymptotic distribution, so that the reported RMSE and power account for estimation uncertainty in the structural model as well as randomization variance.

[25] The appendix also reports results for the demand-only outcome (expected school quality without supply adjustment). That estimand is larger, and power is adequate for most designs, so the bias–power trade-off is less informative as a design tool.

FIGURE 7.—Scale-Up RCT Design



(a) Santiago: $r = 2.0$, $I = 0.7$ km (b) Simulation results

Notes: Panel (a): one stratified random assignment of hexagonal clusters to treatment (teal) and control (gray) in Santiago. Dark inner hexagons are the core zone; lighter outer zones are the treatment ring. Panel (b): RMSE and power for each (r, I) configuration. Highlighted cell minimizes RMSE subject to power $\geq 80\%$. School-level regression on ΔQ_j for voucher and private schools, hex-clustered SEs, $\tau = 0.25$, and no buffer strip ($B = 0$).

Figure 7 presents the results. Panel (a) shows one realization of the hexagonal design overlaid on Santiago ($r = 2.0$ km, $I = 0.7$ km). Teal hexagons are treated clusters, gray are control; darker inner hexagons mark the core zone. Panel (b) shows RMSE and power heatmaps over the design space, with the constrained optimum highlighted.

Results. The two panels of Figure 7(b) reveal a clear bias–power trade-off. RMSE decreases monotonically with hex radius r: larger clusters place core schools farther from hex boundaries, reducing contamination from cross-boundary demand spillovers. Power increases with core radius I: a wider core includes more schools in the regression, shrinking standard errors. The tension is that the lowest-RMSE configurations (large r, small I) tend to have low power, because a tiny core inside a large hex leaves few schools for estimation.

The constrained optimum—the design that minimizes RMSE subject to power $\geq 80\%$—is $r = 4.0$ km, $I = 1.5$ km (RMSE $= 0.017$, power $= 97\%$). More conservative choices such as $r = 2.0$ km, $I = 1.0$ km achieve similar power (96%) at somewhat higher RMSE (0.026). Across the grid, no design with adequate power has RMSE exceeding 0.04, indicating that the equilibrium quality effect ($\beta^* \approx 0.05$ s.d.) is detectable under a range of practical configurations.

The model thus provides concrete guidance for the scale-up design: the locality of competitive interactions—the fact that demand and supply spillovers decay with distance—makes within-market cluster randomization feasible, and the simulation grid identifies the configurations that balance bias against power. A researcher would choose among the well-powered designs based on operational constraints (e.g., how many hexagons are administratively manageable) and the relative weight placed on bias versus precision.

8. DISCUSSION

This paper studies how a low-cost information policy affects school choice and what its effects would be at scale. In the experiment, treated families—especially those reached before they had enrolled—choose schools with higher test scores and value added, and their children score higher on later standardized tests. These results show that a simple intervention delivered through existing preschool infrastructure can shift families toward better schools and improve later academic outcomes.

The policy question, however, is not only whether the intervention works in a small-scale trial, but what happens once it is implemented broadly enough to change market conditions. Our counterfactuals show that this extrapolation is not mechanical. Capacity constraints attenuate the demand-side gains due to congestion when more families seek access to the same high-quality schools. At the same time, the increase in demand for quality changes schools' incentives, and in our preferred specifications the resulting supply-side quality response offsets, and in some cases more than offsets, the attenuation from congestion. The quantitative effects depend on assumptions about assignment, take-up, cost pressure, and which schools adjust. Across these exercises, however, the broader lesson is stable: equilibrium forces are central for evaluating a demand-shifting policy in a school market.

The methodological contribution comes from the complementary roles of the experiment and the model. We do not rely on the structural model alone to infer the demand shift induced by the policy; the randomized trial provides that object directly and credibly. But no small-scale experiment can identify the equilibrium effects that arise only when enough families are treated to change congestion, assignment outcomes, and schools' quality incentives. The model provides that equilibrium extension. It also allows us to study policy margins that are directly relevant for implementation but cannot be learned from the experiment alone, including alternative targeting rules, partial take-up, spillovers across family types, and different assumptions about how schools respond. In that sense, the framework turns a small-scale experiment into a disciplined

tool for prospective policy analysis rather than treating it as a stand-alone estimate of policy impact.

The same framework is useful for experimental design. When a policy is rolled out gradually or across only part of a set of urban markets, the design of the larger-scale evaluation determines what can be learned about equilibrium effects. Our scale-up exercise shows how the estimated model can be used to choose clusters and buffer zones in a way that balances statistical power against contamination from spillovers. More broadly, the paper illustrates how a credible small-scale experiment, combined with enough economic structure to analyze market adjustment, can provide disciplined guidance about both the likely equilibrium effects of a new policy and the design of the larger-scale experiments needed to learn more about it.

REFERENCES

AGTE, PATRICK, CLAUDIA ALLENDE, ADAM KAPOR, CHRISTOPHER A. NEILSON, AND FERNANDO OCHOA (2024): "Search and Biased Beliefs in Education Markets," NBER Working Paper No. 32670. [16]

ANDRABI, TAHIR, JISHNU DAS, AND ASIM IJAZ KHWAJA (2017): "Report cards: The impact of providing school and child test scores on educational markets," *American Economic Review*, 107 (6), 1535–63. [1, 2]

BERRY, STEVEN, JAMES LEVINSOHN, AND ARIEL PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica: Journal of the Econometric Society*, 841–890. [22]

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market," *Journal of Political Economy*, 112 (1), 68–105. [52]

BROWN, ZACH Y. AND JIHYE JEON (2024): "Endogenous Information and Simplifying Insurance Choice," *Econometrica*, 92 (3), 881–911. [2, 13, 16, 17, 28, 46]

CAMPOS, CHRISTOPHER (2024): "Social Interactions, Information, and Preferences for Schools: Experimental Evidence from Los Angeles," NBER Working Paper No. 33010. [1]

CHUMACERO, ROMULO A., DANIEL GOMEZ, AND RICARDO D. PAREDES (2011): "I would walk 500 miles (if it paid): Vouchers and School Choice in Chile," *Economics of Education Review*, 30 (5), 1103 – 1114. [1]

CORRADINI, VIOLA (2024): "Information and Access in School Choice Systems: Evidence from New York City," Working Paper. [1]

DINERSTEIN, MICHAEL AND TROY SMITH (2015): "Quantifying the Supply Response of Private Schools to Public Policies," Tech. rep., Stanford Institute for Economic Policy Research. [1, 3]

GALLEGO, FRANCISCO AND ANDRES E. HERNANDO (2008): "On the Determinants and Implications of School Choice: Semi-Structural Simulations for Chile," *Journal of LACEA Economia*. [1]

HANSEN, LARS PETER (1982): "Large sample properties of generalized method of moments estimators," *Econometrica: Journal of the Econometric Society*, 1029–1054. [51, 52]

HASTINGS, JUSTINE S. AND JEFFREY M. WEINSTEIN (2008): "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics*, 123 (4), 1373–1414. [1]

HECKMAN, JAMES J, LANCE LOCHNER, AND CHRISTOPHER TABER (1998): "General-Equilibrium Treatment Effects: A Study of Tuition Policy," *American Economic Review*, 88 (2), 381–386. [1]

KAPOR, ADAM, CHRISTOPHER A. NEILSON, AND SETH D. ZIMMERMAN (2020): "Heterogeneous Beliefs and School Choice Mechanisms," *American Economic Review*, 110 (5), 1274–1315. [1]

LISE, JEREMY, SHANNON SEITZ, AND JEFFREY SMITH (2015): "Evaluating search and matching models using experimental data," *IZA Journal of Labor Economics*, 4 (1), 16. [3]

MIGUEL, EDWARD AND MICHAEL KREMER (2004): "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72 (1), 159–217. [37]

MURALIDHARAN, KARTHIK AND VENKATESH SUNDARARAMAN (2015): "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India," *The Quarterly Journal of Economics*, 130 (3), 1011–1066. [37]

NEILSON, CHRISTOPHER A. (2026): "Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students," *Econometrica*, conditionally Accepted. [1, 3, 6, 9, 10, 11, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 27, 30, 31, 35, 46, 47, 48, 51]

OECD (2016): *Starting strong IV: Monitoring quality in early childhood education and care - Chile*, OECD Publishing. [6]

PETRIN, AMIL (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110 (4), pp. 705–729. [22]

SANCHEZ, CRISTIAN (2017): "Understanding School Competition Under Voucher Regimes," Job market paper, phd dissertation, ITAM. [19, 36]

TINCANI, MICHELA M. (2020): "Teacher Labor Markets, School Vouchers and Student Cognitive Achievement: Evidence from Chile," Working Paper. [33]

TODD, PETRA E. AND KENNETH I. WOLPIN (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96 (5), 1384–1417. [3]

——— (2023): "The Best of Both Worlds: Combining Randomized Controlled Trials with Structural Modeling," *Journal of Economic Literature*, 61 (1), 41–85. [3]

TRAIN, KENNETH E. (2009): *Discrete Choice Methods with Simulation*, Cambridge University Press, 2nd ed. [17]

WALTERS, CHRISTOPHER R. (2018): "The Demand for Effective Charter Schools," *Journal of Political Economy*, 126 (6), 2179–2223. [1, 3]

WOLLMANN, THOMAS G. (2018): "Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 108 (6), 1364–1406. [35]

APPENDIX

A1. *Reduced-Form Regression Specifications*

Tables 1 and 2 report intent-to-treat coefficients from reduced-form regressions of the form

$$Y_i = \alpha + \beta T_p + Z_p'\delta + X_i'\gamma + \mu_c + \varepsilon_i,$$

where T_p is treatment assignment at the preschool-center level, μ_c are preschool-municipality fixed effects, and standard errors are clustered at the preschool-center level.

In the main specification, the randomization controls are nine market-level variables: the average SIMCE score used in the report card, the number of schools in the market, the 10th, 25th, 50th, 75th, and 90th percentiles of the SIMCE distribution, the interquartile range of SIMCE scores, and the standard deviation of SIMCE scores. The family controls are indicators for the mother having completed high school and the mother having completed higher education, together with a missing-data indicator for mother's education. This narrow set is the specification used in the paper and is aligned with the family heterogeneity dimension in the structural model.

Panel A uses the full experimental sample. Panel B restricts to families who had already enrolled at the time of the Pre-K visit, and Panel C restricts to families who had not yet enrolled. For Table 1, column 1 restricts the distance outcome to observations with distance below 4 km. Columns 2–6 use the subsample with non-missing school value added so that the school-choice outcomes are compared on a common sample. For Table 2, columns 7–8 are conditional on taking the college entrance exam, and the reported scores pool the 2023–2025 cohorts standardized within cohort.

A2. *Experiment Moments for Structural Estimation*

Table 3 reports the experimental moments used in the second-step demand estimation. These regressions use the model-overlap sample described in Section 5.5.3: families who had not yet enrolled at the time of treatment, can be matched to the six markets covered by the baseline demand model, pass the geographic consistency filter, and have non-missing school characteristics used in estimation.

The regression specification uses the same controls as in Appendix A1: preschool-municipality fixed effects, clustering at the preschool-center level, the nine market-level SIMCE controls, and the mother-education indicators with missing-data flag. The table reports the pooled estimates

together with estimates separately for SEP and non-SEP families. The five moments entering the GMM objective are quality (normalized value added Q_2), sticker price, out-of-pocket price, Google Maps driving distance restricted to observations within 4 km, and voucher-school attendance.

A3. *Scaled-Logit Derivation*

This subsection provides the formal derivation behind the scaled-logit interpretation summarized in Section 5.3. The purpose is interpretive: it shows why imperfect information about school characteristics can motivate both characteristic-specific treatment effects and a common scale shift in the choice model. The demand estimation and counterfactual exercises in the main text are implemented directly in terms of the effective parameters, so none of the policy results requires separately identifying the underlying signal process.

Suppose that for each characteristic $c \in \{q, \text{op}, d\}$, family i of type k observes a noisy signal

$$\tilde{s}_{ij}^{c} = c_j + v_{ij}^{c},$$

where $v_{ij}^{c} \sim N(0, \sigma_{v,c,k}^2)$ is independent noise. Given a prior $c_j \sim N(0, \sigma_c^2)$, the posterior mean is

$$E[c_j \mid \tilde{s}_{ij}^{c}] = \rho_k^{c} \cdot \tilde{s}_{ij}^{c}, \qquad \rho_k^{c} = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_{v,c,k}^2} \in (0,1), \tag{15}$$

where ρ_k^{c} is the signal-to-noise ratio for characteristic c and type k.

Substituting these posterior means into the indirect utility from Equation 1 yields expected utility of the form

$$U_{ij}^{\mathcal{E}} = \phi_k^{q} q_j - \phi_k^{op} \,\text{op}_{k(i),j} + \phi_k^{d} d_{\text{loc}(i),j} + \bar{\beta}\, x_j + \xi_j + \widetilde{\epsilon}_{ij}, \tag{16}$$

where $\phi_k^{q} = \beta_k \rho_k^{q}$, $\phi_k^{op} = \alpha_k \rho_k^{op}$, and $\phi_k^{d} = \lambda_k \rho_k^{d}$ are products of structural preferences and signal precision. Thus, noisier signals attenuate the effective weight placed on the corresponding characteristic. The composite error is

$$\widetilde{\epsilon}_{ij} = \phi_k^{q} v_{ij}^{q} - \phi_k^{op} v_{ij}^{op} + \phi_k^{d} v_{ij}^{d} + \epsilon_{ij}. \tag{17}$$

The variance of this composite error is

$$\text{Var}(\widetilde{\epsilon}_{ij} \mid k) = (\phi_k^q)^2 \sigma_{v,q,k}^2 + (\phi_k^{op})^2 \sigma_{v,op,k}^2 + (\phi_k^d)^2 \sigma_{v,d,k}^2 + \frac{\pi^2}{6}. \tag{18}$$

As in Brown and Jeon (2024), improving signal precision can therefore affect not only the effective coefficients on informed characteristics but also the overall error variance of the choice model.

We approximate this environment with a scaled logit:

$$\widetilde{\epsilon}_{ij} \approx \mu_k \cdot \epsilon_{ij}^*,$$

where ϵ_{ij}^* is i.i.d. EV1 and the type-specific scale is chosen to match the variance of the composite error:

$$\mu_k = \sqrt{1 + \frac{6\,\sigma_{\text{noise},k}^2}{\pi^2}} \geq 1, \tag{19}$$

with

$$\sigma_{\text{noise},k}^2 \equiv (\phi_k^q)^2 \sigma_{v,q,k}^2 + (\phi_k^{op})^2 \sigma_{v,op,k}^2 + (\phi_k^d)^2 \sigma_{v,d,k}^2.$$

Dividing expected utility by μ_k then yields standard logit choice probabilities with effective parameters equal to structural parameters divided by the error scale.

The baseline demand estimation in Neilson (2026) identifies these effective parameters under the normalization $\mu_k^C = 1$. If treatment changes signal precision, the treated and control scales need not coincide. The common scale parameter in Section 5.2 is therefore

$$\kappa_k \equiv \frac{\mu_k^C}{\mu_k^T}. \tag{20}$$

When treatment improves signal precision, $\mu_k^T < \mu_k^C$ and hence $\kappa_k > 1$: choice behavior becomes sharper even for school attributes about which the intervention provides no direct information. When $\kappa_k = 1$, the model collapses to the standard specification with only characteristic-specific treatment effects. In the main text we treat this signal model as one interpretation that motivates allowing a common scale shift; the paper's identification and counterfactual analysis

proceed directly in terms of the effective parameters and do not require separately estimating the latent signal process.

A4. *Estimation Specifics*

In the first step, we estimate the baseline demand parameters (θ_1, θ_2) using the MPEC formulation inherited from Neilson (2026). This approach treats the vector of mean utilities δ as an optimization variable and exploits the sparsity of the Jacobian of the market-share equations. Let $m_{\text{micro}}(\delta, \theta_2)$ denote the vector of micro moments from school choice and let $m_{\text{IV}}(\delta, \theta_1, \theta_2)$ denote the vector of IV orthogonality conditions for the demand residual. The first-step problem can then be written as

$$(\theta_1^*, \theta_2^*, \delta^*) = \arg\min_{\theta_1, \theta_2, \delta} \begin{bmatrix} m_{\text{micro}}(\delta, \theta_2) - \bar{m}_{\text{micro}} \\ m_{\text{IV}}(\delta, \theta_1, \theta_2) \end{bmatrix}' \begin{bmatrix} W_{MM} & 0 \\ 0 & W_{IV-D} \end{bmatrix} \begin{bmatrix} m_{\text{micro}}(\delta, \theta_2) - \bar{m}_{\text{micro}} \\ m_{\text{IV}}(\delta, \theta_1, \theta_2) \end{bmatrix}. \tag{21}$$

The constraints are:

$$\begin{aligned} \delta &= s^{-1}(\bar{S}, \theta_2) && \text{(BLP inversion matching observed shares)} \\ \xi_{jt}(\delta, \theta_1, \theta_2) &= \delta_{jt} - f_{jt}(\theta_1) && \text{(demand residual)} \\ \xi^{\text{norm}} &= 0 && \text{(normalization restrictions),} \end{aligned}$$

where $f_{jt}(\theta_1) = \sum_r \eta_k^r x_{jt}^r$ is the linear index in observed school characteristics. The normalization restriction means that, within each market-year, mean utilities are defined relative to one designated normalization school, so δ_{jt} and ξ_{jt} are identified only up to a market-year-specific additive constant. Thus, $m_{\text{micro}}(\delta, \theta_2)$ corresponds to model-implied averages of the chosen school characteristics by family type, $\bar{m}_{\text{micro}}$ is the corresponding vector in the data, and $m_{\text{IV}}(\delta, \theta_1, \theta_2)$ stacks the sample analog of $\mathbb{E}[Z_{jt}\xi_{jt}] = 0$ for the excluded demand instruments Z_{jt}.

A4.1. *Demand Moments and Excluded Instruments*

The baseline demand system is inherited from Neilson (2026). For convenience, we summarize here the moments and excluded instruments used in that first-step estimation, since they are central to the interpretation of our treatment and equilibrium exercises.

Moments. The first-step demand estimation combines three types of moments. First, *market share moments* match observed school-level shares and recover mean utilities through the BLP inversion. Second, *micro moments* match the average quality, out-of-pocket payment, and distance chosen by each family type in the micro-data. Third, *IV moments* impose orthogonality between the demand residual and excluded instruments. These moments play different roles. The market-share moments are what recover the school-level unobserved demand component, the micro moments discipline heterogeneity in preferences across family types, and the IV moments address the endogeneity of price and academic quality. Thus the micro moments complement, rather than replace, the market-share inversion and IV conditions. In the baseline demand system inherited from Neilson (2026), school-level choice data are available very broadly, while family-specific distance information is available for a smaller but still very large geocoded subsample; accordingly, the distance micro moments are constructed from that subsample, whereas the market-share and IV moments use the broader administrative panel. For our experimental second step, the distance moment is likewise built from family-specific home-to-chosen-school driving distance, while the counterfactual simulations use node-to-school driving distances for the full choice set once each family is mapped to a residential location (node) within a market.

Why instruments are needed. Price and academic quality are endogenous school-side choice variables: schools may choose them in response to unobserved demand components. Following Neilson (2026), let the estimated demand residual be $\widetilde{\xi}_{jt} = \xi_j + \Delta\xi_{jt}$, where ξ_j is a fixed non-academic school attribute and $\Delta\xi_{jt}$ is a period-specific shock. The key maintained conditions are that ξ_j is orthogonal to the instruments and that $\Delta\xi_{jt}$ is independent of both the instruments and the school's price and quality choices. Under these assumptions, excluded instruments shift school incentives without directly moving household demand through unobserved school appeal.

Excluded demand instruments. The main excluded instruments are:

- *Policy-exposure instrument:* the timing of the SEP reform interacted with a school's exposure to SEP-eligible students in its surrounding neighborhood.
- *Simulated transfer instrument:* the change in government transfers implied by the SEP reform when school composition is held fixed at its pre-policy level.
- *Local labor-cost shifter:* a location-by-year proxy for the cost of hiring skilled workers, constructed from wage residuals outside the school sector.

The first two instruments primarily shift schools' revenue and pricing incentives. The local labor-cost shifter is the most direct excluded source of variation for academic quality, because quality provision in this setting is teacher-intensive and labor spending accounts for the large majority of school costs. Together, these instruments provide excluded variation for the endogenous price and quality choices in the baseline demand system.

In the second step, we estimate the treatment-effect parameter vector θ_2^T by matching treatment effects in the experimental data to treatment effects simulated by the model. Let

$$m_{\text{RCT}}(\theta_2^T)$$

denote the vector of second-step moment residuals. In the specification with separate scale parameters for SEP and non-SEP families, this vector stacks the twelve moments described in Section 5:

- treatment effects on chosen quality, sticker price, out-of-pocket payment, and distance, each by SEP status;
- treatment effects on voucher-school attendance, by SEP status; and
- treatment effects on residualized mean utility $\tilde{\delta}_j \equiv \delta_j - \bar{\beta}^q q_j$, by SEP status.

Each element of $m_{\text{RCT}}(\theta_2^T)$ is therefore a difference between an experimental treatment coefficient estimated in the data, $\hat{\beta}^{\text{RCT}}$, and the analogous coefficient computed from simulated choices, $\hat{\beta}^{\text{sim}}(\theta_2^T)$. The second-step problem is

$$\theta_2^{T*} = \arg\min_{\theta_2^T} m_{\text{RCT}}(\theta_2^T)' W_{RCT} m_{\text{RCT}}(\theta_2^T). \tag{22}$$

In the common-κ specification, the $\tilde{\delta}$ moments are omitted and $m_{\text{RCT}}(\theta_2^T)$ contains ten moments instead of twelve.

In the third step, we estimate supply-side parameters (θ_3). To do so, we first construct a residual that measures the gap between observed quality incentives and the fitted quality first-order condition.

When we rearrange the first order condition for quality we can get an expression for the unobserved component that affects the marginal cost of raising quality:

$$\omega_{jt} = \underbrace{\frac{v_b + p_{jt} - \sum_l \gamma^l w_{jt}^l}{\left[q_{jt}^* + s_{jt}(\mathbf{q}, \mathbf{p}, \boldsymbol{\xi}) \left[\frac{\partial s_{jt}(\mathbf{q}, \mathbf{p}, \boldsymbol{\xi})}{\partial q_{jt}}\right]^{-1}\right]}}_{A_{jt}} - \gamma_q \tag{23}$$

Where $\omega_{jt} = \overline{\omega}_j + \Delta\omega_{jt}$, with $\overline{\omega}_j$ defined as the time-invariant component so that $E[\Delta\omega_{jt} \mid j] = 0$ by construction. We exploit the panel structure of the data to concentrate out the school fixed effect $\overline{\omega}_j$. Let the right-hand side of Equation 23 be $A_{jt}(\theta_3)$. Averaging within school gives

$$\overline{A}_j(\theta_3) = \overline{\omega}_j + \gamma_q,$$

so the within-school residual

$$\Omega_{jt}(\theta_3) \equiv A_{jt}(\theta_3) - \overline{A}_j(\theta_3)$$

captures the time-varying component of the marginal-cost disturbance. In practice, the market-year effects γ_{mt} are concentrated out jointly with $\overline{\omega}_j$ via the iterative procedure described in Section 5.5.5. These are the "cost disturbance" moments referred to in the appendix and in the referee comments.

Let Z_{jt}^S denote the excluded supply instruments. The supply moments are the sample analog of

$$\mathbb{E}[Z_{jt}^S \Omega_{jt}(\theta_3)] = 0,$$

and we stack them in a vector $m_{\mathrm{S}}(\theta_3)$. In the implementation used here, Z_{jt}^S is built from the SEP indicator and its interactions with school and market characteristics, so identification comes from how the staggered rollout of SEP shifts the quality first-order condition across schools and over time. The third-step GMM problem is

$$\theta_3^* = \arg\min_{\theta_3} m_{\mathrm{S}}(\theta_3)' W_{IV-S} m_{\mathrm{S}}(\theta_3). \tag{24}$$

A4.2. *Computational Details*

Numerical integration. The demand model includes a single random coefficient on quality ($\nu_i \sim N(0, \sigma^2)$). All integration over this random coefficient—in the market share formula (Equation 2), in the simulated treatment effects (Equation 12), and in the supply-side markdowns—uses 5-point Gauss-Hermite quadrature, a deterministic rule that is exact for polynomials up to degree 9. No Monte Carlo simulation draws are used in the estimation.

Weighting matrices. The first-step demand weighting matrix $W = \text{diag}(W_{MM}, W_{IV\text{-}D})$ is inherited from Neilson (2026). The second-step (treatment effects) and third-step (supply) estimations each use two-step efficient GMM: the first GMM step uses the identity matrix $W^{(1)} = (1/n_m) \cdot I$, and the second uses the inverse of the estimated covariance matrix of the moments: $W^{(2)}_{\text{RCT}} = \hat{V}^{-1}_{\text{CR}}$ (clustering at the preschool center level) for Step 2, and $W^{(2)}_{\text{S}} = \hat{V}^{-1}_{S}$ for Step 3.

Optimization algorithms and initialization. Step 1 inherits the published baseline demand estimates from Neilson (2026); our code loads those estimates rather than re-estimating them in this paper. Step 2 uses a constrained optimization algorithm with analytical gradients, a non-negativity bound on the SEP quality-shift parameter, bounds $\kappa \in [0.5, 4.0]$, and a multi-start strategy. In the preferred two-κ specification reported in the paper, we use 7 starting points in each of the two GMM steps; all 7 converge to the same objective value and the same parameter estimates in the identity-weighted step, and all 7 converge to the same objective value and the same parameter estimates in the efficient-GMM step. The optimization tolerances are 10^{-8} (optimality) and 10^{-10} (step size). Step 3 concentrates out the school-specific costs ω_j and market-year effects γ_{mt} via an alternating-projection algorithm, reducing the free parameter space to γ_{SEP}. The concentrated objective is minimized using a Nelder-Mead simplex algorithm with tolerances of 10^{-12} (function) and 10^{-10} (parameter).

A5. *Calculating Standard Errors*

The standard errors of the estimated parameters in each step of the estimation procedure are obtained from the variance-covariance matrix for the GMM estimator proposed by Hansen (1982). We will discuss how we calculate the standard error for a generic case (the parameters θ and moments $\mathbf{M}$), and then we discuss the case of each set of parameters more specifically.

Each one of our GMM estimators is the result of an optimization problem in which the objective function has a quadratic form:

$$\min_{\theta} Q_{obj} = \mathbf{M}' W_m \mathbf{M}$$

For which the gradient is:

$$\frac{\partial Q_{obj}}{\partial \theta} = 2 \cdot \mathbf{J}_M' W_m \mathbf{M}$$

Where $\mathbf{J}_M$ is the Jacobian.

The variance-covariance matrix for a GMM estimator is calculated using the estimator proposed by Hansen (1982):

$$\text{cov}(\theta) = (\mathbf{J}_M' W_m \mathbf{J}_M)^{-1} \mathbf{J}_M' W_m \mathbf{V} W_m' \mathbf{J}_M (\mathbf{J}_M' W_m \mathbf{J}_M)^{-1}$$

Where V is the variance-covariance matrix of the moments. With deterministic Gauss-Hermite quadrature, there is no simulation error in the model predictions. For the parameters estimated from the experimental moments, the relevant source of uncertainty is therefore sampling error in the treatment-effect estimates. As discussed by Berry et al. (2004), when both sources are present they enter additively; in our current implementation only the sampling component remains. The RCT moments in Equation 12 take the difference between the estimated treatment effects and the model's predictions, so the variance of those moments is estimated from the cluster-robust variance-covariance matrix of the underlying treatment-effect regressions.