

NBER WORKING PAPER SERIES

EXPERIMENTALLY VALIDATING WELFARE EVALUATION OF SCHOOL VOUCHERS

Peter Arcidiacono
Karthik Muralidharan
John D. Singleton

Working Paper 32968
<http://www.nber.org/papers/w32968>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2024, Revised September 2025

None of the authors have any financial interest in the results reported in this paper. We thank Matteo Bobba, Michael Dinerstein, Chao Fu, Adam Kapor, Anoosheh Mirkushal, Christopher Neilson, Isaac Sorkin, Petra Todd, and seminar participants at Barcelona, Cowles, the Federal Reserve Bank of Minneapolis, Johns Hopkins, Michigan, Michigan State, NBER Education, Princeton, Stanford, Toulouse, UPenn, Western Ontario, and Yale for helpful comments and discussions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Peter Arcidiacono, Karthik Muralidharan, and John D. Singleton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Experimentally Validating Welfare Evaluation of School Vouchers
Peter Arcidiacono, Karthik Muralidharan, and John D. Singleton
NBER Working Paper No. 32968
September 2024, Revised September 2025
JEL No. H4, I20, O12

ABSTRACT

We leverage a unique two-stage experiment that randomized access to private school vouchers in rural India to estimate willingness-to-pay for greater school choice, a key input for assessing welfare impacts of alternative voucher designs. We find that a voucher targeted to households with limited assets, a proxy for ability-to-pay constraints, yields a marginal value of public funds (MVPF) greater than 3. We obtain this result using a multistep research design that isolates the information contributed by experimental data. We begin by estimating several models of school choice on control data alone, reserving data from treatment markets for model selection via out-of-sample validation. All of the models underpredict take-up of the randomized voucher offer. Evidence from treatment markets points to why: the program induced household search and led private schools to use voucher surplus to attract students. We incorporate these mechanisms into a unified model, estimated with both treatment and control data, that successfully explains the observed choice patterns. Our experience highlights that credible policy analysis requires modeling choice frictions and supply responses and using experimental data in estimation.

Peter Arcidiacono
Duke University
Department of Economics
and NBER
psarcidi@econ.duke.edu

John D. Singleton
University of Rochester
Department of Economics
and NBER
john.singleton@rochester.edu

Karthik Muralidharan
University of California, San Diego
Department of Economics
and NBER
kamurali@ucsd.edu

1 Introduction

Governments around the world provide in-kind benefits such as public schooling, health care, and food assistance (Currie and Gahvari, 2008). A central question in public economics is whether in-kind provision is less or more efficient than giving beneficiaries vouchers to purchase the same goods or services on the open market. A large empirical literature has studied this question across sectors and contexts, and most studies evaluate vouchers by their impact on sector-specific outcomes such as test scores, food consumption, or nutrition. This likely reflects the priorities of paternalistic policymakers seeking to assess impacts on outcomes of interest.

An alternative approach to evaluation, rooted in welfare economics, is to ask how vouchers affect well-being. Conceptually, this is measured by beneficiaries’ willingness to choose the voucher over the in-kind option, which also reflects valuation of outcomes not measured by a researcher. Estimating voucher valuation is also essential for predicting take-up under different voucher values and targeting rules, a key input for policy design. Yet, such valuations are rarely estimated, partly because translating observed choices into welfare requires a structural econometric model. Such models—whether used to integrate under a demand curve or to forecast counterfactual policy impacts—rely on distributional assumptions and exclusion restrictions that are often hard to test.

One promising way forward is to combine structural modeling with datasets featuring random treatment assignment (Attansio and Blair, 2018; Galiani and Pantano, 2023; Todd and Wolpin, 2023). However, there remains debate over how best to do so. One approach is to use experimental data for model estimation, which can aid identification of key parameters. Alternatively, it can be reserved for model validation. This can mitigate concerns with “structural data mining” while providing a basis for model selection—ability to forecast out-of-sample program impacts—that is the criterion that matters for ex ante policy analyses (Schorfheide and Wolpin, 2012, 2016).¹

In this paper, we study both approaches, and thereby illustrate how experimental data can improve the credibility of structural econometric models in practice. Our context is estimating households’ willingness-to-pay for private school vouchers in rural India. We first ask whether school choice models fit to observational control data validate against the experimentally-induced take-up patterns of a randomized voucher trial. Finding that the models *do not validate*, we then diagnose sources of mis-specification and use those insights to develop an improved model. This model, estimated with the experimental data, is used to evaluate welfare impacts of voucher

¹Well-known studies using Mexico’s PROGRESA program illustrate both approaches: Attanasio, Meghir and Santiago (2012) use experimental data for model estimation, while Todd and Wolpin (2006) use it for validation.

programs under alternative targeting criteria.

Our design leverages data from the Andhra Pradesh School Choice Project, a randomized school voucher trial.² The project’s unique market-level randomization of villages let us reserve experimental data for model validation, simulating the information set facing a researcher conducting ex ante policy analysis.³ We do this by embargoing all data from treatment markets during the development and estimation of several structural models of school choice (described below). For validation, we assess the models’ accuracy at predicting school choices in treatment markets, including of those who were randomly offered a private school voucher. This exercise maintains the assumptions that the models are identified on the control market data alone and that the voucher affected choices only via its direct reduction of tuition and fees.⁴ To enhance the design’s credibility, we pre-committed to the models, their predictions, and to specification tests in a publicly-circulated working paper, Arcidiacono et al. (2021), before lifting the treatment data embargo.

We estimate two classes of choice models on the control markets data in which households choose a primary school from among the free government and fee-charging private options in their village. The first are random coefficient logit demand models that are standard for welfare analysis (Berry, Levinsohn and Pakes, 1995; Nevo, 2001; Petrin, 2002), and have been applied to other contexts of school choice (e.g. Neilson 2013; Carneiro, Das and Reis 2022). Second, we develop and estimate a random utility discrete choice model that incorporates a latent constraint on households’ ability-to-pay for private schooling in the absence of a voucher, reflecting the reality that liquidity and credit frictions plausibly constrain choices in this setting.⁵ Our constrained-choice framework parallels other applications where choice sets are not observed (e.g. Ben-Akiva and Boccara 1995; Barseghyan et al. 2021) and connects with work quantifying the salience of credit constraints.⁶

To use the treatment data for model selection via validation, variation in control markets sufficient to identify the price (tuition and fees) elasticity of demand for private schools is essential. The observational data from control markets include information on each sampled child’s school and demographics, characteristics of the government and private school options in each village, and geocoded school and household locations used to calculate travel distances—i.e. “micro data”

²The trial’s test score impacts are reported in Muralidharan and Sundararaman (2015).

³Randomization of applicants in treatment markets to voucher status followed the randomization of villages into control and treatment markets.

⁴This latter assumption is reasonable given the project’s limited scale and one-time implementation.

⁵For instance, Tarozzi et al. (2014) find that micro consumer-loans substantially raised ownership and use of insecticide-treated bednets in rural India, while demand was highly elastic when households had to pay upfront. In our data, 41% of households whose child attends a government school cite “economic reasons” for their choice.

⁶Examples include Cameron and Heckman (2001); Keane and Wolpin (2001); Gregory (2017); Delavande and Zafar (2019).

(Berry and Haile, 2024). We combine the spatial variation in school access with standard supply-side instruments—e.g. location in product space and cost measures (e.g. Berry, Levinsohn and Pakes 1995; Hausman 1996; Nevo 2001)—to identify preferences. To separate willingness- from ability-to-pay, we assume that observed household assets, which relate to credit access, are conditionally unrelated to differences in indirect utilities. Both the instruments and the constraint specification could well be mis-specified; we estimate and generate predictions for several specifications of each model type, highlighting the prospective value of using out-of-sample experimental variation for testing modeling and identification assumptions.

We find that both classes of models fit the data from control markets well, but the ability-to-pay constrained model delivers markedly different estimates. In particular, the constrained model ascribes greater average utility from private schools and to attributes such as English-language instruction (offered only by some private schools). This is especially so for low asset households, about a quarter of whom are estimated to have no feasible private school option without a voucher. This difference has substantial welfare implications: our ability-to-pay constrained model suggests that households who choose a private school because of the voucher value the offer *more* than the average private school household.

How do the models estimated using only control markets data perform out-of-sample in the treatment data? Among voucher applicants in control markets, 27% attended private schools. The random coefficient model predicts a 28 point increase in private schooling (to 55%), whereas the ability-to-pay constrained model, consistent with the greater value assigned to private schooling, predicts a 38 point increase (to 65%). Yet, the actual increase was 58 points, with 85% of voucher winners attending private schools, and we can statistically reject equality between predicted and actual take-up. Thus, though our ability-to-pay constrained model fits relatively better, it still substantially underpredicts experimental take-up of the voucher offer.

We use the data from treatment markets to understand why the models fail to validate. We discover two facts that contradict the assumption underpinning validation that voucher offers only affected choices by setting tuition to zero for recipients. First, while the models match choice patterns of households who were not eligible for the voucher, or who did not apply, private school attendance among voucher applicants who are lottery *losers* is nearly 50% higher than among comparable applicants in control markets. The voucher program thus affected choices of voucher losers in some way. Second, the models especially under-predict the rate at which voucher winners attend *low tuition* private schools. This is counter-intuitive since the value of free-tuition should,

all else equal, be higher in high-tuition schools. This pattern points away from classic issues with instruments and towards behavioral responses by low tuition schools that evidently increased their appeal to voucher winners.

Based on the data, we propose two mechanisms to reconcile these findings. First, the possibility of a voucher appears to have induced greater school search in treatment villages, even among voucher losers. Voucher losers' private school attendance is elevated even in several treatment markets where the voucher program was effectively absent.⁷ This suggests that seeing some applicants win vouchers may have led losers to also consider private schools for their children.⁸ Second, private schools had strong incentives to recruit voucher students, especially in low-fee schools where the voucher value (which was set at the 90th percentile of the private school fee distribution) was substantially higher than marginal cost. Evidence from a post-intervention survey shows that siblings of voucher winners were more likely to receive scholarships, suggesting that schools passed through some of the surplus to households to attract voucher winners. Alternative mechanisms such as peer effects or advertising by private schools are not supported in the data.

Since the treatment induced behavioral changes among both households and schools that were absent in control markets, we develop a unified choice model incorporating these mechanisms into the ability-to-pay constrained model. Search is modeled as a cost households incur to reveal match quality at private schools in their village. We assume that voucher losers in treatment villages expect to receive a voucher, but otherwise face the same post-search choice environment as control households. Those that draw a sufficiently high match quality may therefore choose to attend a private school even if they have to pay tuition and fees. To capture enrollment incentives, we include the voucher surplus—the difference between the voucher amount and a private school's tuition and fees—in utility (if positive). These two mechanisms add three parameters to the model.

We estimate the unified model on pooled control and treatment data. The household-level randomization of offers within treated markets, interacted with variation in voucher surplus across private schools, identifies enrollment incentives, while identification of search frictions leverages the villages where the program wasn't implemented. The unified model does not guarantee a match between actual and predicted voucher usage; there are no treatment group dummies in the model. Thus, the fact that the unified model successfully explains choice patterns in control markets,

⁷This was the case in 9 markets for reasons explained in Section 5.1.

⁸This could reflect a combination of losers being told that there was a randomized waiting list from which additional vouchers may be offered, and losers accompanying friends or neighbors who were winners to learn more about private schools. The search mechanism is also consistent with evidence from other settings that suggest information about schools is costly to acquire. See e.g. Arteaga et al. (2022); Larroucau et al. (2024); Neal and Root (2024).

the elevated private school attendance among voucher losers in treatment markets, and whether and where voucher winners take-up the offer increases confidence in the estimates. The unified model estimates show that some of the unobserved heterogeneity attributed to liquidity and credit constraints in the control data was due to search frictions, but ability-to-pay constraints remain empirically relevant.

We use estimates from the unified model for welfare analyses under counterfactual policies. Our base case is a universal voucher of the same value as in the experiment, but capped at the tuition of the school. This policy generates gains in consumer surplus for compliers who switch from government to private school, and also generates fiscal savings since the per-student cost of government schools is over 2.5 times the voucher value. However, a universal voucher also incurs extra costs by covering students who would have attended private schools anyway. Put together, we estimate that such a program would have a marginal value of public funds (MVPF) of 1.33-3.05. Targeting vouchers (of the same value) only to asset-poor households generates larger gains: this is self-financing when the fiscal savings are large (implying an MVPF of infinity) and has an MVPF of 3 even under conservative assumptions on fiscal savings. Finally, targeting vouchers specifically to ability-to-pay constrained households—as estimated by the unified model—would more than double the MVPF.

This paper makes both substantive and methodological contributions. Substantively, we show that voucher programs can meaningfully raise economic well-being, as measured by willingness-to-pay, when private provision is more efficient or when vouchers are well-targeted, as illustrated above. These findings complement prior work on voucher programs, including on the AP School Choice Program (Muralidharan and Sundararaman, 2015), which has mainly focused on test-score impacts (Epple, Romano and Urquiola, 2017).⁹

Our focus on welfare, as revealed by choices, connects with a growing literature that measures preferences over schools and school attributes (e.g., Hastings, Kane and Staiger, 2005; Bayer, Ferreira and McMillan, 2007; Neilson, 2013; Abdulkadiroğlu et al., 2020), including applications to valuing school choice and voucher programs (e.g., Ferreyra, 2007; Carneiro, Das and Reis, 2022; Kamat and Norris, 2025). Highlighting the importance of viewing schools as differentiated products, we find that households induced to attend private school by a universal voucher would pay up to 25% of median per capita annual consumption for the program—far more than for a one

⁹Evidence from international settings tends to show positive effects, including in the longer run (e.g. Angrist, Bettinger and Kremer 2006).

test score standard deviation gain in school value-added. Moreover, we show that school choice is shaped by meaningful frictions. While limited information has also been emphasized in other recent work (e.g., Andrabi, Das and Khwaja, 2017; Agte et al., 2024; Corradini, 2023), we explicitly model the role of credit and liquidity constraints and find them to be important.

How we use the RCT data offers lessons for combining structural econometric models with randomized experiments, as an expanding body of work does (e.g., Duflo, Hanna and Ryan, 2012; Galiani, Murphy and Pantano, 2015; Lagakos, Mobarak and Waugh, 2023). We set out along the rare path of using experimental data only for model selection via validation, including taking steps to disincentivize over-fitting and blind model development to treatment patterns.¹⁰ A recognized requirement for this approach is that the models be identified using only the control data (Todd and Wolpin, 2023). But our exercise revealed a distinct limitation: the models failed because the voucher program induced household and school responses unanticipated at the outset. This journey highlights two broader lessons: First, structural models that neglect choice frictions and supply-side responses are unlikely to be reliable for out-of-sample policy analyses, even of limited-scale interventions. Second, our experience underscores advantages of using experimental data to fit—as opposed to validate—structural models: the treatment data both brought to light missing mechanisms and proved necessary to identify parameters governing them.

2 Background and Research Design

Our data are drawn from a randomized controlled trial conducted in 180 villages in the Indian state of Andhra Pradesh (AP). Villages selected for the project had to have at least one private school that agreed to participate in the voucher program. As a result, more than one of every two primary school students (57%) across the project villages attended a private school at baseline (2008).¹¹

The program was targeted to students likely to otherwise attend government schools. Students randomized into treatment status were offered a voucher covering the costs of tuition and required expenses (e.g. books and uniforms) at government-recognized, participating private schools in their village for the duration of primary schooling. At the average private school in the project, tuition

¹⁰Our decisions to embargo treatment markets data and to pre-commit to predictions are similar to Pathak and Shi (2021), which validates structural school choice models fit prior to a policy change in Boston. However, two distinguishing elements of our setting are: 1) endogenous tuition and fees charged by private schools; and 2) unobserved heterogeneity in households’ choice sets arising from ability-to-pay constraints (and search costs).

¹¹Based on household survey data from AP project districts (ASER, 2018). The private school market share *not* conditioning on whether the village has at least one private school was around 33% in 2008 (ASER, 2018). The study was conducted from 2008-12 and is set in the erstwhile undivided state of Andhra Pradesh, which was later divided into the two states of Andhra Pradesh and Telangana in 2014.

and fees were otherwise about Rs. 1,900 per year (Table A2)—equivalent to nearly 8% of median annual consumption per capita.¹² The annual voucher value was set at around the 90th percentile of the tuition and fee distribution among private schools (Rs. 2,600) and was paid directly to schools. Expenses for transportation, however, were not covered by the voucher and, unlike government schools, private schools in this setting do not provide free mid-day meals.

Participation in the project at the school-level was voluntary, but participating private schools were not allowed to screen or selectively admit voucher students. Government-recognition was required to participate.¹³ The design stipulated that lotteries would be held to allocate places in oversubscribed schools, but in practice this proved unneeded.

2.1 Research Design

An important feature of the AP School Choice project is its two-stage randomization: At baseline, parents of eligible students were invited to apply for the program with the knowledge that the voucher would be allocated by lottery and that applying would not guarantee receipt. After eliciting interest from eligible households, the project first randomized villages into 90 treatment and 90 control markets. Applicant households in treatment villages were then randomized into (“voucher winners”) or out of (“voucher losers”) of receiving a voucher offer.¹⁴

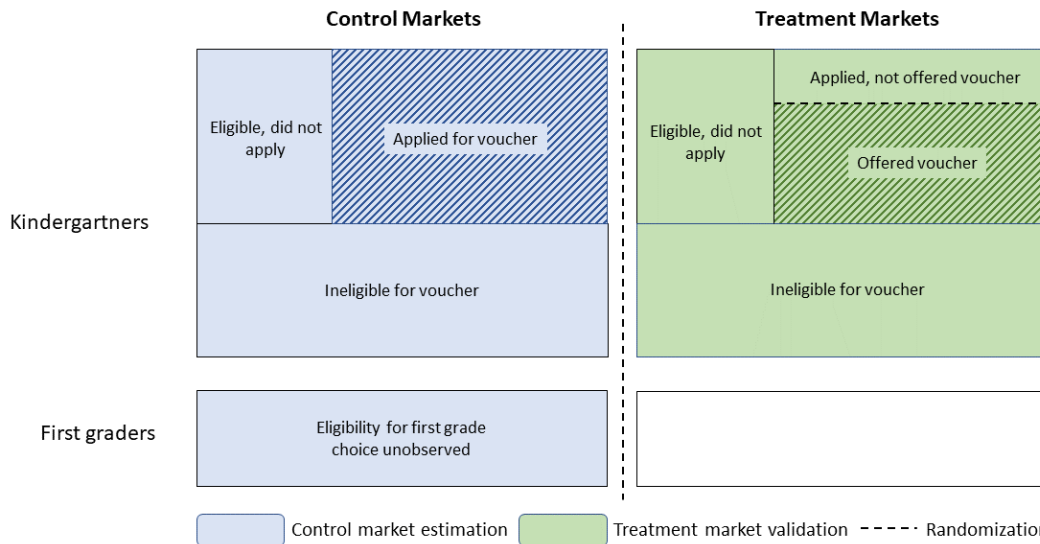
Our research design leverages the market-level randomization to first estimate school choice models on data from control markets before validating the models against choice patterns in treatment markets. The AP project experiment was conducted in parallel on two cohorts of students: a younger cohort, who had yet to enter primary schooling at time of baseline and who we term kindergartners, and an older cohort of first graders already attending a primary school at baseline. Our empirical models focus on households’ first grade primary school choice. This has two consequences for our research design: 1) in fitting the models, we use kindergartners’ choices subsequent to baseline and first graders’ *retrospective* choices¹⁵; and 2) the experimental validation by necessity

¹²Median household expenditure per capita was about Rs. 24,000 per the 2011-12 India Human Development survey in comparable rural villages (with a private school) of Andhra Pradesh.

¹³Note that several unrecognized private schools that sought government recognition failed to secure this before the program launched, contributing to the fact that almost no students were able to successfully use a voucher in a handful of villages. We use attendance patterns in these “non-complying” villages to later test for and separate mechanisms influencing choice patterns.

¹⁴This double randomization design facilitated estimating spillover effects on non-participants in the program, and also provided a pure control group (in the control villages) that would be unaffected by such spillovers. See Muralidharan and Sundararaman (2015) for details.

¹⁵We use retrospective choices because first graders’ choices after baseline—whether to remain in their current primary school or *switch*—are different in nature from the choice of where to begin primary school due to switching costs, changes in information, and differences in the voucher offer’s present value.



Notes: Figure visually represents the paper’s research design, which 1) estimates models of primary school choice on first grade choices of kindergartners and first graders in control markets; and 2) validates the models using data from treatment markets. Eligibility for AP voucher (for first grade choice) determined by attending Anganwadi (government-run pre-school). Vertical dashed line represents first stage market-level randomization; horizontal dashed line dividing treatment market applicants represents household-level randomization. Shading with upward-sloping diagonal lines represents experimental validation sample.

Figure 1: Control Market Sample and Treatment Market Validation for First Grade Choice

focuses on the choices of kindergartner voucher winners.

Figure 1 presents a visual representation of the design. Shown by the light blue shading, several empirical models of primary school choice, detailed in the next section, are fit to only the data from control markets. The control markets data are purely observational; we observe school choices made by households, characteristics of those households, and attributes of the school options (including the tuition and fees). As the figure shows, the kindergarten control markets sample contains several subgroups: those who were ineligible for the voucher; those who were eligible and did not apply; and those who were eligible and applied but were randomized-out at the village-level. Eligibility was determined by whether the child attended a government preschool (Anganwadi) or not, which was used as a proxy means test, since students from better off households were more likely to be attending a private preschool.

The out-of-sample validation step of our research design uses the data from treatment markets, which include parallel subgroups except that the household-level randomization further splits kindergartner applicants into voucher winners and voucher losers. For kindergartners in treatment villages that did not receive a voucher (e.g. those ineligible), we evaluate how the models fit out-of-sample. For those that received a voucher offer, we evaluate the models based on their predictions for choice patterns—e.g. what share would take-up the voucher offer (and where do they tend to

use it). These exercises rely on the assumption that the voucher influences choices in treatment markets only by defraying voucher winners’ tuition and fees. We do this by counterfactually setting tuition and fees at government-recognized private schools to zero in the models. This experimental validation step is visually represented by the boxes overlaid with upward-sloping diagonal lines.

An important clarification about our experimental validation is that the “take-up” number reported in Muralidharan and Sundararaman (2015) is not directly comparable to the end point outcome we use. In particular, their tabulation excludes several groups of students who, under the logic of our counterfactual simulations, would be classified as having accepted the voucher. This includes students who were later deemed ineligible—such as those who independently gained admission to a private school—as well as voucher students who initially enrolled in private school but subsequently switched to a government school. Appendix C provides details on how we define and code take-up for the purposes of validation in this paper.

2.2 Sample Construction and Summaries

In keeping with our research design, we construct household-level control and treatment markets samples in separate (identical) steps. These samples record the school choice of every household while linking them to a village-specific set of possible alternatives. The samples were constructed from a baseline household survey, which elicited rich information on the demographic and socioeconomic characteristics of students and households, and a survey of schools that began in the first year of the program. GPS coordinates of households and schools were collected, facilitating the mapping of travel distances. Students were assessed at baseline and during the third and fourth years of the program. We build the lists of schooling options by taking the union of primary schools in the baseline household survey and those that respond to the school survey; the average village includes over 6 private school options. We restrict the samples to students who attend a school in their village.¹⁶ The control and treatment market samples contain 4,251 and 4,123 unique households, respectively. In treatment villages, 629 kindergartner students were randomly offered a voucher.

Table A1 summarizes characteristics of households across all project villages (control and treatment). Stratifying first graders by their school choice reveals that private school students are more likely to have parents who both completed primary school; more likely to have a parent who completed secondary school; and more likely to live in a pucca (brick or stone) house, have a water

¹⁶One issue this raises is attrition of kindergartners after baseline data collection, which further may be affected by the voucher offer in treatment markets. We address this in estimation by weighting households’ likelihood contribution accordingly.

facility in the home, and to have a household toilet. The average private school students also scores at least three fifths of standard deviation higher at baseline in math and the local language, Telugu. In addition, the table shows that (kindergartner) students eligible for the program are more similar in background demographics and socioeconomic status to (first grader) government school students than to private school students.

Summarizing the information collected from the school survey, Table A2 compares government and private schools across all markets, highlighting differences that accompany the fact that private schools are fee-charging.¹⁷ A larger share of government schools are pucca structures and nearly all include a library, whereas a greater share of private schools have a functioning toilet and a staffroom for teachers. A much greater share of teachers in government schools have a bachelor’s degree or a formal teacher training credential. Private school teachers are more likely to be female, from the village, and present in the classroom during enumerator visits. Private schools also have less multi-class teaching. The majority of private schools feature instruction in English. Moreover, many private schools teach Hindi and computer skills (whereas few government schools do). Though not shown in the table, private schools also have a longer school year (2 weeks additional instructional time) and a longer school day (45 minutes more per day).¹⁸

The randomization and symmetric data collection (and processing) imply baseline balance on average between 1) control and treatment market schools; 2) control and treatment market households who did not (or could not) apply for the voucher; and 3) control and treatment market households who applied for the voucher.¹⁹ Consistent with this, Table A1 shows limited statistical differences between control and treatment market subgroups of households. Table A2 likewise shows school-level balance along most dimensions. However, treatment market private schools are about 13% more expensive and this difference is robust to controls.

3 Control Models and Results

In this section, we describe our empirical models of household school choice that were estimated using only data from the control villages. In the choice models, we treat households, which consist

¹⁷For cases where the school is not in the initial survey year, we fill their characteristics with information from the household survey or taken from subsequent school survey years whenever possible. In particular, we use later survey years to impute missing tuition and fees based on the school’s percentile in the overall distribution when observed.

¹⁸Further, there are significant differences in allocation of instruction time with private schools spending less time on math and Telugu and using the extra time to teach additional subjects (see Muralidharan and Sundararaman 2015 for details).

¹⁹There is also balance between treatment market applicant households randomly offered and randomly not offered a voucher.

of at least one primary school aged child, as unitary decision makers. As private schools charge tuition and fees, households must weigh the expected benefits of private school attendance against foregone consumption. Such benefits potentially include a more attractive combination of school amenities as well as human capital gains.

Given the presence of liquidity and credit frictions in our setting, we explicitly model the influence of an unobserved ability-to-pay constraint on choice of primary school. In relaxing this constraint, a private school voucher thereby potentially generates welfare benefits by expanding households' choice sets, potentially allowing them to select their most preferred option. In addition, we compare the predictions of this class of model, which for identification places structure on how observed measures of household wealth influence choices, with those obtained from random coefficient demand models. Differences in income may map to differences in preferences in the random coefficient model, but all households are assumed able to choose from any primary school in their village.

3.1 Ability-to-Pay Constrained Choice

In selecting a primary school, households weigh the utility of the school alternatives that belong to their village.²⁰ This set is denoted by \mathcal{V}_i for household i . However, the tuition and fees may exceed the household's ability-to-pay. This is captured in the model through a constraint on their choice problem:

$$\max_{j \in \mathcal{V}_i} U_{ij} \geq U_{ij'} \quad \forall j' \in \mathcal{V}_i \text{ where } p_j, p_{j'} \leq \omega_i \quad (1)$$

For any school, j , the household's consumption and tuition and fees, denoted p_j , must not exceed the household's ability-to-pay, which we denote by ω_i . For government schools, p_j is zero (or nearly so). The ability-to-pay constraint represents the combination of a household's income and accumulated savings with their ability to borrow against future income to finance private schooling. This "reduced-form" constraint on choice also captures the possibility of subsistence or liquidity constraints.

Households rank the available schooling alternatives according to an indirect utility function. Letting α represent household i 's marginal utility of consumption, the indirect utility to household i of school choice j can be written as:

$$U_{ij} = \alpha(y_i - p_j) + X_j' \beta_i + \gamma_i \ln D_{ij} + \delta \text{Closest}_{ij} + \xi_j + \epsilon_{ij} \quad (2)$$

²⁰Primary schooling is compulsory, so we do not model the choice of whether to send the child to school or not.

D_{ij} is the distance between school j and household i 's home, while X_j represents school characteristics. $Closest_{ij}$ captures the possibility that the nearest government school is especially salient. In estimation, we include in X_j whether a school is government or private, is government recognized (if private), is English medium, offers Hindi classes, is connected to a secondary school, and three indices capturing, respectively, the quality of facilities, teachers, and teacher characteristics. Also contained in X_j is school j 's value-added in math, which we estimate from the panel of student test scores using standard methods.²¹ ξ_j represents an index of commonly-valued amenities of school j unobserved to the econometrician and likely correlated with tuition. ϵ_{ij} is assumed to follow a Type 1 extreme value distribution.

We subscript the parameters in equation (2) by i to denote their dependence on observed household characteristics, W_i :

$$\begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \gamma_1 \end{pmatrix} + \begin{pmatrix} \beta_2 \\ \gamma_2 \end{pmatrix} W_i$$

The household characteristics in W_i mediate the valuation households place on school amenities, capturing systematic heterogeneity across households in willingness-to-pay. W_i includes AP voucher program eligibility status and indicators for gender, whether belongs to a scheduled caste, is Muslim, whether an older sibling attends government school, whether both parents completed primary school, and whether one parent completed secondary school.²²

3.1.1 Instrumenting for Private School Tuition and Fees

A first empirical challenge for identifying this model (which applies equally to the random coefficient class discussed next) on the control markets data is that ξ_j is unobserved. We implement a control function approach to address the endogeneity of private school tuition and fees (Petrin and Train, 2010).²³ This strategy regresses tuition and fees on school characteristics and a set of instruments in a first stage:

$$p_j = X_j' \Gamma + f(Z_j) + \mu_j \quad (3)$$

²¹Appendix A details the value-added estimation. We also include indicators for missing distance, missing value-added, and imputation of tuition and fees in the utility function.

²²Our specifications do not include all possible interactions of household and school characteristics. The exact interactions included are summarized in Table A3.

²³This is because the AP project's sampling design means that not all schools are attended by a household in our estimation sample, so shares cannot be inverted.

where X_j are observed school characteristics (including the estimated value-added), Z_j are instruments, and $E[\xi_j \mu_j] > 0$. The utility function we then ultimately take to the data is given by:

$$U_{ij} = -\alpha p_j + X_j' \beta_i + \gamma_i \ln D_{ij} + \delta \text{Closest}_{ij} + \kappa \hat{\mu}_j + e_j + \epsilon_{ij}$$

where $\hat{\mu}_j$ is the first stage residual for private school j and e_j is a normally-distributed random effect; both terms are zero for government schools.²⁴

Our baseline specification uses two instrumental variables. First, we use a summary measure of each private school’s location in “product space” (Berry, Levinsohn and Pakes, 1995). We do this using factor analysis applied to totals of characteristics of *other* schools in the village for each private school, e.g. the number of other English-medium schools. The second instrument uses the spatial environment to isolate exogenous cost differences across private schools (Hausman, 1996; Nevo, 2001). We construct a proxy for costs for each private school based on the average tuition chosen by similar private schools that are located in *other* villages. In implementation, we match private schools within medium of instruction and focus on other schools not in nearby villages.²⁵ First stage estimates are presented in Table A4.

3.1.2 Identifying and Estimating Ability-to-Pay

The second empirical challenge for estimating the choice problem described by equation (1) is that households’ ability-to-pay, ω_i , is not contained in the data. This introduces unobserved heterogeneity across households in choice sets. Mis-specifying households’ choice of school as unconstrained is liable to bias estimates of willingness-to-pay and underestimate the gains of a voucher.

We specify latent ability-to-pay as a function of observed household wealth factors, given by:

$$\ln \omega_i = I_i' \lambda + v_i \tag{4}$$

In this equation, the household’s log ability-to-pay at the time of choosing a primary school depends on the wealth factors, I_i , and unobservable household-specific v_i . We assume that v is distributed normally, with variance σ , and independent of the choice shocks.

²⁴Beyond assumptions about instrument validity and independence of μ_j , the control function first stage is also inconsistent with a Nash-Bertrand price-setting equilibrium where private schools know the ξ s of competing schools.

²⁵This is to minimize the confounding influence of spatially-correlated demand shocks. As an alternative to the predicted tuition instrument, we also estimate models using the product space IV and a village-level cost index constructed from private schools’ reported expenditure. Arcidiacono et al. (2021) provides additional details on construction of the instruments.

The above structure implies choice set probabilities, which are useful to examine for discussing identification. The basic insight is to recognize that each household i can fall into one of a finite, village-specific number of possible choice sets. Let j_i^* index schools in i 's village in terms of ascending tuition and fees (such that J_i^* is the most expensive). Then denote by $\phi_{ij_i^*}$ the probability that household i is in choice state of being able to afford at most: $p_{j_i^*} \leq \omega_i \leq p_{j_i^*+1}$. We can write this as:

$$\phi_{ij_i^*} = \Phi\left(\frac{\ln p_{j_i^*+1} - I_i' \lambda}{\sigma}\right) - \Phi\left(\frac{\ln p_{j_i^*} - I_i' \lambda}{\sigma}\right)$$

where the state probability is a difference between points on the normal CDF that depend on data (tuitions and I_i) and parameters (λ and σ). $\Phi\left(\frac{\ln p_1 - I_i' \lambda}{\sigma}\right)$ is the probability of not being able to choose *any* private school in their village.²⁶ Combining logit expressions for choice probabilities with the choice set probabilities allows us to form a likelihood for each household.

The choice probability expressions highlight the data elements that influence ability-to-pay. In particular, to separate from willingness-to-pay, it is key that I_i and W_i do not overlap completely. We assume that assets (such as home ownership) can be excluded from utility. The asset variables thus only enter I_i , being used exclusively to proxy for household wealth and ability to borrow (Filmer and Pritchett, 2001). In practice, we use the first principal factor obtained from the six household assets in our data. Intuitively, this restriction ascribes differences in choice patterns between households that are similar in non-asset characteristics (e.g. demographics and parental education) but who differ in assets to ability-to-pay differences. Our baseline specification allows only AP voucher eligibility and household size to belong to both W_i and I_i .²⁷

3.2 Random Coefficient

We compare the latent ability-to-pay model with random coefficient models similar to classic demand estimation applications (e.g. Berry, Levinsohn and Pakes 1995; Nevo 2001; Petrin 2002) and the models of school choice in Neilson (2013) and Carneiro, Das and Reis (2022).²⁸ In this class

²⁶ And the probability the household can choose from *all* private schools is given by $1 - \Phi\left(\frac{\ln p_{J_i^*} - I_i' \lambda}{\sigma}\right)$.

²⁷ We estimate alternative specifications on the control data that relax these exclusion restrictions by allowing ability-to-pay to also depend on caste status and letting whether the parents are laborers (a variable not included in the baseline specifications) influence preferences and ability-to-pay.

²⁸ Arcidiacono et al. (2021) also describes and presents predictions for a model that assumes all choices are available and groups households into clusters based on observables, allowing preferences to be cluster-specific. We do not discuss this ‘‘clustered multinomial logit’’ demand model here given its predictions and estimates are qualitatively the same as the random coefficient model.

of models, the underlying choice problem is instead unconstrained—households are able to choose from *any* primary school in their village.

The indirect utility in the random coefficient model is given by:

$$\begin{aligned} U_{ij} &= -\alpha_i p_j + X_j' \beta_i + \gamma_i \ln D_{ij} + \delta \text{Closest}_{ij} + \xi_j + \epsilon_{ij} \\ &= -\alpha_i p_j + X_j' \beta_i + \gamma_i \ln D_{ij} + \delta \text{Closest}_{ij} + \kappa \hat{\mu}_j + e_j + \epsilon_{ij} \end{aligned} \quad (5)$$

where the substitution reflects the control function strategy for addressing unobserved ξ_j , which is applied in the same way. While similar to the ability-to-pay constrained model, this indirect utility differs in two ways: First, note that the function allows for heterogeneity across households in their sensitivity to higher tuition and fees, reflected in the indexing by i . Specifically, we allow α_i to depend on household asset levels (e.g. whether the household owns up to six possible assets) and household size.

Second, the random coefficient model accommodates greater flexibility in how households value school characteristics. Like the ability-to-pay constrained model, the random coefficient model specifies a parametric relationship between observed household characteristics, W_i , and preferences over non-tuition school amenities. However, the random coefficient model includes an additional stochastic component on household preferences for private schooling. Letting β_i^P indicate the marginal utility to household i of attending private school, this parameter can be expressed as:

$$\beta_i^P = \beta_1^P + \beta_2^P W_i + \nu_i \quad (6)$$

ν_i is an unobserved, continuous type that follows a mean-zero normal distribution. This additional stochastic term captures unobserved heterogeneity in preferences for private schooling across households

3.3 Control Markets Estimation and Results

Per our research design, we estimate the empirical models above using only data from the control markets. In addition to baseline specifications of the models, we estimate several alternative specifications that modify the instrument sets, sets of interactions, and weights.

The estimation pools choices from several subgroups of students, shown with the light blue shading in Figure 1: kindergartners who were eligible (by virtue of attending an Anganwadi at

baseline) and who applied for the voucher program; eligible kindergartners who did not apply; ineligible kindergartners; and first graders. Though the model specifications allow for preferences (and ability-to-pay, in the constrained model case) to depend on AP voucher eligibility, we do not model application status.²⁹ However, since eligibility status is unknown for the older cohort, we model latent eligibility of these students and use the EM algorithm for estimation. We treat the private school random effects, which adjust the variance of the private school choice shocks, as iid school- and household- specific and weight each household’s likelihood contribution to account for attrition and for the sampling design of the AP project.³⁰

We report selected estimates of the baseline control model specifications in Table A5 and summarize here what they imply for preferences and welfare.³¹ Estimates from the ability-to-pay constrained model imply that around 13% of applicant households (and 24% of low asset households) are unable to choose *any* private school in their village (absent a voucher). The ability-to-pay constrained control model accordingly ascribes greater utility from private schooling—and from school attributes associated with private schools—than the random coefficient model. Notably, this difference translates into important differences in willingness-to-pay for a voucher, especially for households induced to attend a private school (i.e. compliers). Fit to the same control data, the ability-to-pay constrained and random coefficient models basically agree on the average always takers’s valuation, but the constrained model suggests the average complier’s valuation is larger (while the random coefficient model indicates they value it about half as much).³²

4 Treatment Markets Validation

We now turn to evaluating the empirical models’ out-of-sample performance using the held-out treatment markets data. The treatment data allow for non-experimental and experimental validation. These can each be understood visually from Figure 1. In the treatment data, we have several subgroups of kindergarten households who did not receive a voucher offer: those who were eligible and applied, but randomized out at the household level; those eligible who did not apply; and the ineligible. We can therefore ask how well the models estimated on the control data do in explaining the choice patterns of these households.

²⁹As justification for this, conditional on observables, application status is not a statistically significant predictor of private school attendance in the control data.

³⁰See Appendix B for further detail.

³¹The full set of control model estimates pre-committed to are reported in Tables A24 and A25 of Arcidiacono et al. (2021). The paper also contains a longer discussion of the control model results.

³²Note this is despite the ability-to-pay constrained model also predicting greater compliance (i.e. take-up).

The primary focus of our design, however, is on validation out-of-sample against choice patterns under the voucher experiment. This experimental validation is represented by the boxes in Figure 1 filled with diagonal lines: using the empirical models, we generate predictions for the voucher take-up of kindergartner applicants. We do this by setting tuition and fees at participating private schools to zero and simulating choices. This allows us to compare model-based “treatment” moments with analogous moments calculated directly from the treatment group.

4.1 Non-Experimental Validation

We begin by examining how well the empirical models fit the choice patterns of treatment market households who do not receive a voucher. To do so, we take the cleaned data from treatment markets for ineligible households, eligible non-applicants, and applicants who did not win a voucher and directly apply the control model estimates. This allows us to compute predictions for private school attendance.³³

Table 1: Out-of-Sample Validation: Treatment Market Predictions

	N	Attend Private Model			Tuition Private Model		
		Data	RC	CC	Data	RC	CC
Ineligible for voucher	413	0.99	0.99	0.98	1.87	1.96	2.02
Eligible non-applicants	134	0.16	0.19	0.17	1.95	2.00	2.04
Voucher losers	299	0.43	0.29	0.28	2.12	1.98	2.00
Voucher winners	574	0.85	0.58	0.67	2.09	2.46	2.48

Notes: Table reports attendance at private schools (government-recognized and unrecognized) and average tuition given private attendance among treatment market kindergartner subgroups in the treatment market data (Data) and as predicted by the control random coefficient model (RC) and control ability-to-pay constrained model (CC). Only winners in complying villages are included in bottom row tabulations and model numbers reflect adjustment for group’s selective attrition (Appendix D).

Table 1 shows the results of validating the models out-of-sample. Both the random coefficient model and the ability-to-pay constrained model match well the private school attendance rates for ineligibles and eligible households who didn’t apply for the voucher program. However, both models significantly underpredict private school attendance of voucher losers by nearly 15 percentage points.

To further examine the fit of our models to these groups, we formulate the question of misspecification as hypotheses tests in the spirit of tests for “forecast bias” (e.g. Chetty, Friedman and Rockoff 2014). To do so, we begin by fixing the indirect utility for each option j in treatment models

³³This requires that we use the first stage for tuition and fee endogeneity estimated on the control data to compute residuals for treatment market private schools. Consistent with the treatment-control difference in private school tuition, shown in Table A2, this imputes a higher average unobserved quality among treatment market private schools.

to that predicted from the control model estimation (plus a T1EV choice shock). For empirical model m :

$$\hat{u}_{ij}^m = -\hat{\alpha}_i^m p_j + X_j' \hat{\beta}_i^m + \hat{\gamma}_i^m \ln D_{ij} + \hat{\delta} \text{Closest}_{ij} + \hat{\xi}_j^m$$

We then estimate an auxiliary model for each control model on kindergartners in treatment villages who do not win a voucher where their indirect utility at j (less an idiosyncratic choice shock) is specified as:

$$U_{ij}^m = \hat{u}_{ij}^m + \pi_T^m \text{Private}_j + \pi_L^m \mathbf{1}[\text{VoucherLoser}_i] \times \text{Private}_j + \tau^m p_j + \epsilon_{ij}$$

This specification allows us to see whether the overall private utility from the control model (which is embedded in \hat{u}_{ij}^m) is different for treatment village ineligible and eligible non-applicants (given by π_T^m) and for voucher losers (given by $\pi_T^m + \pi_L^m$) as well as whether the price coefficient is different in treatment villages. Under the assumption the model is true, \hat{u}_{ij}^m controls for all observed and unobserved qualities of school j . These auxiliary models thus ask whether the control models do a good job predicting which private school these students attend (as a function of tuition) as well as whether they attend private school.

Table 2: Non-Experimental Validation: Hypotheses Tests

	RC		CC	
	(1)	(2)	(3)	(4)
Private school		-0.12 (0.50)		0.07 (0.09)
Private school \times Voucher loser		2.21 (0.56)		2.05 (0.30)
Tuition and fees (1000s of Rs.)		0.00 (0.10)		-0.11 (0.09)
AIC	2,399	2,260	2,411	2,265

Notes: Table reports hypothesis tests of model mis-specification that examine predictive power of private voucher school constant and tuition and fees for choices of treatment market kindergartners not offered a voucher conditional on the indirect utility of the alternative implied by the control random coefficient model estimates (RC) and control ability-to-pay constrained model estimates (CC). $N = 846$ kindergartner treatment market households not offered a voucher. Excluded group is ineligible and eligible kindergartners who did not apply for AP voucher. Standard errors reported in parentheses.

The results are presented in Table 2. Columns (1) and (3) report goodness-of-fit summaries in the form of AIC stats for the random coefficient and ability-to-pay constrained control models, respectively. The fit of the random coefficient model to the choices of treatment market households who do not win a voucher is marginally better. Columns (2) and (4) show that, consistent with

Table 1, the estimates of the private dummy are not different from zero for ineligible or for non-applicants, but the voucher loser interactions are significantly positive irrespective of the control model. At the same time, the coefficient on price is small and insignificant, suggesting that the control estimates are providing good estimates of the tuition gradient for this sample.

Overall, the non-experimental validation suggests that both control models provide a good fit for the treatment market data for households who did not apply for a voucher or were ineligible for one, but do not fit well for applicants randomized out from receiving a voucher offer. The fact that both models miss for this group raises the question of whether, and in what way, the voucher intervention may have influenced their choices—a question we return to later.

4.2 Experimental Validation

This subsection presents the findings from the experimental validation of the control models. We focus on predictions for voucher take-up before using the hypothesis testing framework to statistically quantify the results and investigate sources of model mis-specification.

4.2.1 Predicted versus Actual Voucher Take-up

We first examine how well the control models predict private school attendance of voucher winners.³⁴ Table 3 presents actual and model-predicted take-up. The first column reports attendance at private schools by applicant kindergartners in control markets. Overall, 27% of applicants in control markets choose to attend a government-recognized private school. Columns (3) and (4) report model predictions for voucher take-up according to the random coefficient and ability-to-pay constrained model, respectively.³⁵

The random coefficient model predicts that private school attendance under the voucher will increase by 23 points to 50%. The ability-to-pay constrained model predicts that private school attendance will more than double, increasing another 10 points to 60% of those offered. Across households, this gap is pretty uniform, though it is only six points among those where both parents completed primary school. Among households where a parent completed secondary school, the

³⁴Appendix C discusses nuances of coding take-up consistent with the counterfactual simulations, which explains why our overall take-up rate is higher than that reported in Muralidharan and Sundararaman (2015).

³⁵These predictions do not exactly match those reported in Arcidiacono et al. (2021) because Table 3 re-computes them on the students offered the voucher in the treatment markets, whereas the original predictions were computed for applicants in control markets. These predictions are thus adjusted for minor treatment-control differences in observables. They also account for non-participation in the program by some schools. The predictions also differ a little because Arcidiacono et al. (2021) simulated take-up allowing households to use a voucher at any private school—not only government-recognized ones, as stipulated by the program.

Table 3: Experimental Validation: Take-up of Voucher Offer

	Data		\hat{U}_{se}		\hat{U}_{se} adj.	
	Control	Treat	RC	CC	RC	CC
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	0.27	0.85	0.50	0.60	0.56	0.65
Female	0.24	0.86	0.50	0.59	0.55	0.64
Muslim	0.47	0.98	0.70	0.79	0.80	0.86
Lower caste	0.18	0.77	0.42	0.53	0.47	0.57
Older sibling in gov't school	0.14	0.79	0.33	0.43	0.40	0.49
Both parents completed primary school	0.41	0.88	0.64	0.70	0.69	0.74
≥ 1 parent completed secondary	0.46	0.76	0.67	0.74	0.71	0.77
Both parents laborers	0.21	0.77	0.44	0.54	0.49	0.59
Asset level < 3	0.21	0.85	0.47	0.57	0.54	0.63
Asset level = 3	0.29	0.85	0.51	0.61	0.56	0.65
Asset level = 4	0.25	0.85	0.50	0.60	0.56	0.65
Asset level > 4	0.38	0.89	0.59	0.66	0.64	0.70

Notes: Table presents average (government-recognized) private school attendance by applicants in control markets (Control), average voucher take-up by treatment market applicants (Treat), and average voucher take-up by treatment market applicants as predicted by random coefficient (RC) and ability-to-pay constrained control models (CC) by subgroup. Columns (5) and (6) adjust the predictions upward for the reduction in winners' attrition due to the voucher offer. Predictions correspond to baseline specification described in the text and detailed in Arcidiacono et al. (2021).

ability-to-pay constrained model underpredicts by only 2 points. Columns (5) and (6) of Table 3 adjust the model predictions for the voucher's impact on attrition. This correction, which assumes that all excess non-attriters among voucher winners take-up, raises the predictions to 56% and 65% take-up of the voucher offer, respectively.³⁶

Column (2) of Table 3 reports take-up of the voucher in the treatment markets—what actually happened. As reported earlier, 85% of applicants randomly offered the voucher attended a private school. Compared with the random coefficient model prediction, this represents a gap of 29 points. The ability-to-pay constrained model's prediction was also too low, but by 9 fewer points. The subgroups comparisons show that the models performed especially badly at predicting take-up of students with an older sibling in government school. The ability-to-pay constrained model was off by 30 points for this group. This reflects that the control market estimates of both models assign a significant disutility to attending a private school for these students. The data-prediction gap in the case of the constrained model is 15 points among households without an older sibling in government schools. While these comparisons discussed pertain only to the baseline specifications of the control models, the gaps highlighted are robust across the alternative control model specifications estimated (e.g. using the alternative IVs).³⁷

³⁶Appendix D discusses how we implement this correction in greater detail.

³⁷Table A7 compares model predictions for *elasticities* of private schooling with respect to the voucher offer with those computed using the experimental variation, revealing similar findings. However, the table reveals that the

Table A8 compares effects of the voucher offer on characteristics of households’ chosen schools in terms of treatment-control differences. As expected, the offer raised the (posted) tuition at chosen schools (by about Rs. 1000 on average) and increased attendance at an English medium school by 13 points. It also increased attendance at a school offering Hindi by 33 points. Both models underpredict the effect on Hindi, but produce similar ITT effects as the experiment on English and tuition despite underpredicting private school attendance significantly.

4.2.2 Hypothesis Tests

We examine model fit and mis-specification by estimating auxiliary models on the choices of the treatment market applicants randomly offered a voucher while controlling for the indirect utility predicted by the control models. Specifically, we estimate models of the following form:

$$U_{ij}^m = \hat{u}_{ij}^m + \hat{\alpha}_i^m p_j + \pi_V^m PrivateVoucher_j + \epsilon_{ij} \quad (7)$$

for each empirical model m estimated on the control sample. $\hat{u}_{ij}^m + \hat{\alpha}_i^m p_j$ is treated household i ’s predicted indirect utility from choice j , according to the estimates from control model m . Like before, if control model m accurately captures treated students’ take-up of the voucher offer (i.e. their preferences over voucher-eligible private schools), we expect that $\pi_V^m = 0$. For households we code as intending to use the voucher but who were not able to actually use it, estimation matches their intended use with their probability of attendance at any government-recognized private school in their village with the voucher.

Columns (1) and (4) of Table 4 report measures of goodness-of-fit to offered students’ choices under the voucher; the constrained model achieves a lower AIC. Columns (2) and (5) then add an intercept for private (government-recognized) schools, as in the hypothesis testing framework outlined above. This provides an alternative way to quantify underprediction of take-up between control models: the coefficient on the intercept is 40% larger per the random coefficient model.

Columns (3) and (6) of Table 4 simultaneously estimate an intercept for voucher-eligible private schools and a “slope” on the tuition at those schools:

$$U_{ij}^m = \hat{u}_{ij}^m + \hat{\alpha}_i^m p_j + \pi_V^m PrivateVoucher_j + \tau_V^m p_j + \epsilon_{ij}$$

difference in the elasticity between the low and high asset households is matched more closely by the ability-to-pay constrained model.

Table 4: Experimental Validation: Hypothesis Tests Comparing Random Coefficient and Ability-to-pay Constrained Models

	(1)	RC (2)	(3)	(4)	CC (5)	(6)
Private school		4.72 (0.30)	7.49 (0.46)		2.60 (0.22)	5.28 (0.40)
Tuition and fees			-1.32 (0.17)			-1.32 (0.16)
AIC	1,496	1,198	1,135	1,400	1,235	1,164

Notes: Table reports hypothesis tests of model mis-specification that examine predictive power of private voucher school constant and tuition and fees for voucher winners’ choices conditional on the indirect utility of the alternative implied by the control random coefficient model estimates (RC) and control ability-to-pay constrained model estimates (CC). Note that both regressors are zeroed out for voucher-ineligible private schools (i.e. those without government recognition) for this analysis. $N = 574$ kindergartner treatment market households offered a voucher (not in non-complying treatment villages). Standard errors reported in parentheses.

The results reveal that, while both models underpredict voucher use, they *over*-predict usage at higher tuition private schools. In other words, offered students use the voucher at lower tuition schools than expected. Further, the coefficient on tuition is remarkably similar between models and, though not shown in the table, this pattern holds across levels of household wealth. This finding is key for understanding the sources of mis-specification in the control market estimates. In particular, it suggests that conventional unobserved school characteristics are not the issue. Rather, if there is an unobservable school “quality” that voucher winners are sorting on, it is negatively correlated with tuition.

4.3 Key Findings

The treatment markets validation reveals several clear empirical patterns. First, the baseline specifications taken to the control markets data of both the random coefficient and ability-to-pay constrained models provide accurate predictions of private school enrollment for households who did not apply for the voucher offer in treatment markets. However, all control models substantially underestimate private school attendance among voucher losers, suggesting that the voucher program influenced these households’ decisions beyond the direct effect of tuition reductions.

Additionally, the experimental validation demonstrates that all of the control models significantly underpredict actual voucher take-up among voucher winners. While the ability-to-pay constrained model performs better, it still falls short of observed take-up levels. Strikingly, voucher users also disproportionately choose lower tuition private schools. This finding contradicts stan-

dard expectations that unobserved school quality and tuition positively correlate, suggesting that an alternative mechanism must be driving these decisions.

In Appendix E, we show that, while additional specification tests suggest some evidence for other sorts of mis-specification of the ability-to-pay constrained model, the central patterns of underprediction are not meaningfully altered by the inclusion of additional interactions.

5 What Was Missing?

The out-of-sample validation revealed significant discrepancies between predicted and observed voucher take-up, as well as unexpected spillovers onto voucher losers. These empirical patterns suggest that our working assumption that the voucher affected school choices only through direct tuition reductions was incorrect. In this section, we turn to analyses of the treatment markets data to investigate potential behavioral mechanisms absent from our control models.

5.1 Voucher Program Increased Private Schooling Among Voucher Losers

That voucher losers in treatment villages enrolled in private schools at much higher rates than their counterparts in control villages suggests the intervention impacted their choices in some way.³⁸ One potential way the treatment could have impacted the behavior of voucher losers is if they too expected to receive a voucher in the future and therefore decided to explore private school options. This could reflect a combination of losers being told that there was a waiting list from which additional vouchers may be offered, and them accompanying winning friends or neighbors to learn more about private schools.

Evidence supporting this explanation comes from nine treatment villages where, in the end, virtually no household was able to use a voucher. This occurred because all the private schools in these villages ended up not accepting the voucher.³⁹ One reason is that private schools needed to be government-recognized to participate and several failed to secure this recognition prior to the program's launch. A second was the testing requirement, which some schools found burdensome. We thus examine whether voucher applicants—both winners (who could not use the voucher, *and* losers—in these non-complying villages have private school attendance rates similar to voucher losers in other treatment villages. To examine this, we estimate a linear probability model where the

³⁸A question this raises is whether the village-level randomization was successful and we confirm that control and treatment market applicants are largely balanced, including on stated schooling preference at baseline.

³⁹Note that *some* private schools in other treatment villages also did not participate in the end, but voucher winners were still able to use the voucher at other private schools in those villages that did participate.

dependent variable is private school attendance. Denoting the non-complying treatment villages as “flagged,” we control for (i) whether the student applied for the voucher interacted with treatment village and flagged villages, (ii) whether they were offered a voucher interacted with flagged villages, (iii) where they were ineligible for the voucher. All coefficients are relative to the omitted group: eligible students who did not apply for a voucher in control villages.

Table 5: Private School Attendance in Non-compliant Treatment Villages

	Attend voucher private	
Offered AP voucher	0.377*** (0.030)	0.412*** (0.031)
Offered × Flagged village		-0.399*** (0.115)
Applied for AP voucher	0.068** (0.033)	0.068** (0.033)
Applied × Treatment village	0.155*** (0.039)	0.154*** (0.039)
Applied × Flagged village		-0.001 (0.115)
Treatment village	-0.029 (0.026)	-0.025 (0.027)
Flagged village		-0.043 (0.061)
Ineligible for AP voucher	0.622*** (0.031)	0.623*** (0.030)
Constant	0.197*** (0.030)	0.197*** (0.030)
Observations	2,960	2,960

Notes: Table reports estimates of linear probability models of private school attendance among kindergartners to examine differences in attendance pattern in “flagged” non-complying treatment villages where the program was not successfully implemented. Excluded group is AP voucher-eligible households who did not apply. Standard errors reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The results are shown in Table 5. Coefficients on treatment village and flagged village are small and significant, consistent with treatment-control balance at the village level and flagged villages being similar to other treatment villages. The coefficient on applied for the voucher indicates that voucher applicants are 6 percentage points more likely to attend private school in control villages than those who are eligible who did not apply, consistent with selection into application. The interaction between applying for a voucher and treatment village is large and positive (and matches the evidence shown earlier of a 15 point discrepancy), as is the effect of winning a voucher.

What is most interesting is what happens in flagged treatment villages. Now we no longer see an effect of winning a voucher: the coefficient on winning a voucher interacted with flagged village is of the same magnitude as that on winning a voucher but of the opposite sign. Yet, both voucher

winner and loser attend private schools in flagged villages at similar rates to voucher losers in other treatment villages, and at higher rates than applicants in control villages: the coefficient on applied for a voucher interacted with flagged village is zero.⁴⁰

Overall, the patterns in Table 5 are consistent with voucher losers and voucher winners in flagged villages equally anticipating a voucher offer. These students then explore private school options about private schools under that pretense. Some of these students then draw sufficiently high match qualities to rationalize elevated private school attendance even after it was later revealed they would have to pay tuition.

While the results are consistent with peer effects in *search*, they are inconsistent with peer-effects in *attendance*. If voucher losers were more likely to attend because of the higher attendance of voucher winners, then we would expect their attendance in flagged villages—where vouchers could not be used—to be much lower. Yet, voucher losers are as likely to attend private schools in these villages as in other treatment villages.⁴¹

5.2 Voucher Winners Disproportionately Choose Low-Tuition Private Schools

We next turn to why our models missed the high attendance rate of voucher winners at low-tuition schools. We propose pass through of voucher surplus as an explanation. Given that the voucher’s yearly value was set to the 90th percentile of the tuition and fee distribution of the private schools in these villages (i.e. around 44% more than the annual tuition and fees charged by the average private school), and that a private school enrolling a voucher student would receive this amount regardless of their own tuition and fees, strong incentives were in place for low tuition schools to attract voucher students. A profit-maximizing private school with tuition below the voucher amount may then try to attract voucher students by sharing the surplus generated, and, importantly, this incentive will be stronger for lower tuition private schools.

We focus on pass through because it can be plausibly targeted to voucher winners, consistent with the data patterns. For example, the voucher program could equally have given low tuition private schools an incentive to increase demand through advertising. But if advertising were an

⁴⁰Here one may be concerned that these flagged villages were not randomly assigned. But note that all of the flagged village coefficients (with the obvious exception of the interaction with offered a voucher) are small and insignificant, suggesting these schools are similar to those in other treatment villages with the exception of the ability to use the voucher.

⁴¹Additional evidence that peer effects cannot explain elevated private school attendance among voucher losers is presented in Figure A1, which relates the variation in the share of voucher-winning Anganwadi peers to the likelihood non-offered applicants attend private school: Rather than a positive association between the share of Anganwadi peers who are offered a voucher and private school attendance, we find essentially no relationship.

explanation, it should affect the decisions of all students. However, in treatment villages, there is no evidence of increased attendance for eligible non-applicants and no evidence of higher attendance at low-tuition schools conditional on attending a private school for either eligible non-applicants or voucher losers.⁴²

Table 6: Voucher Pass-through: Survey Responses for Focal Child and their Siblings

	(1) Private	(2) Tuition and fees (Rs.)	(3)
Offered voucher	0.542*** (0.0277)	-2,742*** (199.5)	-580.5*** (113.1)
Constant	0.220*** (0.0424)	3,153*** (263.1)	760.5*** (127.1)
Observations	948	395	941
Sample	All	Private=1	All
Siblings (ages 5-9)			
Offered voucher	0.152*** (0.0470)	-860.9** (392.4)	289.2 (179.0)
Constant	0.265*** (0.0851)	1,396*** (444.9)	313.6* (181.5)
Observations	452	183	441
Sample	All	Private=1	All

Notes: Table reports ITT estimates of voucher offer impact on private school attendance (column 1) and spending on tuition and fees (column 3) according to post-intervention survey data on focal study child (upper panel) and their primary school-aged siblings (lower panel). Column (2) examines differences in spending on tuition and fees conditional on the focal study child attending a private school. Each upper panel observation is a study kindergartner; each lower panel observation is a school-aged sibling of a study kindergartner. Sample includes kindergarten voucher winners in treatment villages and voucher applicants in control villages for which survey data is available. Standard errors reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

While low tuition schools had incentives to attract voucher students, a challenge for the pass through explanation is how the surplus could feasibly be shared with voucher students. We present evidence for one way this was achieved which is through offering scholarships to voucher students' siblings. Specifically, we examine post-intervention survey responses of households in control and treatment markets regarding private school attendance and their expenditure on tuition and fees. Importantly, the survey includes responses pertaining to the focal child, who may or may not have received a voucher, as well as for their siblings in the household. Focusing on kindergarten voucher winners in treatment villages and voucher applicants in control villages, the top panel of Table 6 shows, as expected, that those randomly offered a voucher report 54 point greater private school attendance for the main child (column 1) and report spending about Rs. 600 less on the main

⁴²Note that advertising targeted specifically at voucher winners would look exactly like enrollment incentives.

child’s tuition and fees (column 3). The bottom panel of the table reports analogous intent-to-treat estimates for school-aged siblings of the main child. The key finding is that the offer raises the probability their sibling attends private school by 15 points (column 1) without changing the household’s spending on tuition and fees for the sibling child (column 3).⁴³

6 Unified Model

Given the results in section 5, we now describe a unified empirical model that incorporates search and pass through (along with an ability-to-pay constraint) which we estimate on the combined data from control and treatment markets. We show the unified model successfully rationalizes the data patterns and finally use it to estimate welfare effects.

6.1 Incorporating Search and Enrollment Incentives

The unified model incorporates two new features:

1. search—households must pay a cost to reveal their match qualities at private schools and all voucher applicants in treatment villages anticipate receiving a voucher; and
2. enrollment incentives—participating private schools in treatment villages share a fraction of the program’s surplus with voucher recipient.

We begin by incorporating enrollment incentives into the indirect utility function. Recall that the school receives V for a voucher student which in almost all cases is higher than the tuition of the school, p_j . For schools where $p_j < V$, we assume that the household and the school Nash bargain over the surplus where the surplus is defined as the difference between the value of the voucher and the tuition and fees of the school: $V - p_j$. Denote θ as the Nash-bargaining parameter for the household, $0 \leq \theta \leq 1$. The ex-post utility (minus the preference shock) from attending private school j with a voucher in our unified model is then given by:

$$u_{ij}^V = u_{ij} + \alpha p_j + \alpha \theta (V - p_j) \times \mathbf{1}[V > p_j] \tag{8}$$

where u_{ij} is the “control” utility previously given by equation (2). αp_j is added to this because these households anticipate receiving a voucher (recall that the coefficient on tuition and fees in

⁴³The middle column of Table 6 (column 2) shows that, conditional on the focal child attending a private school (which is endogenous), offered households report spending essentially zero on the main child’s tuition and fees and report spending about 60% less than control households on their siblings’ tuition and fees.

u_{ij} was α). The strength of the enrollment incentive times the value of the incentive is governed by $\alpha\theta$, the parameter on the difference between the voucher amount ($V = 2.6$ since the program paid private schools Rs. 2,600 for each voucher enrollee) and private school j 's tuition and fees; no incentive is applied if the school's fees exceed the voucher amount. It is intuitive to see from this equation that the enrollment incentive will be larger at low-tuition private schools, potentially reconciling why more voucher winners than expected attend such private schools.

The second extension is to introduce search. In particular, households have full information about government schools, but must pay a cost to reveal their match (represented by the preference shocks, the ϵ 's) with private schools.⁴⁴ Households that cannot afford to attend any private school in their village will then not pay the search cost as, even with the match terms revealed, the private schools are not affordable. But when households have vouchers, participating schools are now in the choice set and households may find it optimal to search.

When households do not search, their choice set is given by G_i , the set of government schools in i 's village; when households search, all schools in the village that the households can afford are in their choice set. Applicants search when the expected utility from choosing among all schools they can afford minus the search cost is higher than the expected utility from choosing only among government schools. These expected utilities in the case of Type 1 extreme value shocks follow directly from McFadden (1978). Letting c_i represent the cost of search, applicants in treatment villages search when:

$$\begin{aligned} c_i &< \ln \left(\sum_{j \in \mathcal{V}_i} \exp u_{ij}^V \right) - \ln \left(\sum_{j \in G_i} \exp u_{ij} \right) \\ &< -\ln(P_{iG|S}^V) \end{aligned} \tag{9}$$

where $P_{iG|S}^V$ is the probability i chooses a government school conditional on searching and receiving a voucher. In contrast, "control" households will search for private schools when

$$\begin{aligned} c_i &< \ln \left(\sum_{j \in \mathcal{V}_i} \mathbf{1}[p_j \leq \omega_i] \exp u_{ij} \right) - \ln \left(\sum_{j \in G_i} \exp u_{ij} \right) \\ &< -\ln(P_{iG|S}) \end{aligned} \tag{10}$$

where, recall, ω_i represents unobserved ability-to-pay. Thus, absent a voucher, constrained house-

⁴⁴In the case of ineligible students, we assume they paid the search cost earlier; the reason they were ineligible for the voucher program is that they were attending a private school pre-kindergarten.

holds will be less likely to pay the search cost because many of the private schools will be outside of their price range regardless, limiting the benefits. With this added mechanism (characterized by two new parameters, the location and scale of c_i which we assume is exponentially distributed), the voucher thus can affect private school attendance both through searching (by increasing the expected gains from search) and by making private schools more attractive conditional on searching.

This search channel can also help explain higher private school attendance by applicants in treatment villages who did not receive a voucher. Consistent with the patterns presented earlier, we treat these households as *expecting* to get the voucher, as in equation (9). Then, at the stage where they must make a decision as to which school to attend, they receive no enrollment incentive and must pay full price at participating private schools (i.e. their ex-post utility is given by u_{ij} , not u_{ij}^V). Applicants in treatment villages see an increase in private school attendance because, by engaging in search, they discover that the match effect was high enough for schools that were affordable to induce them to attend a private school, even though in expectation it would not have been utility-maximizing to search had they known they would not receive a voucher.⁴⁵

6.2 Estimation of the Unified Model

We estimate the unified model on the combined control and treatment markets data, pooling households across all subgroups shown in Figure 1. We focus on extending the ability-to-pay model, both because of its better fit for voucher winners and because search will be less effective in explaining the underprediction for voucher applicants in treatment villages.⁴⁶ As in the control model estimation, we account for the endogeneity of tuition and appropriately weight given sampling considerations. Details of the instrumenting and weighting are included in Appendix G.

We estimate the unified model using a modified EM algorithm where latent eligibility (of first graders) and the structural parameters are estimated in separate maximization steps. Let θ represent the structural parameters underlying search, the ability-to-pay constraint, and utility. At the θ -maximization step, we maximize:

$$\tilde{L} = \sum_i w_i [\tilde{e}_i \ln L_{1i}(\theta) + (1 - \tilde{e}_i) \ln L_{0i}(\theta)]$$

⁴⁵An alternative modeling approach would be to treat voucher losers in treatment villages as having a belief of receiving a voucher that is less than one. We assume that they expect to receive the voucher with probability one based on the results in Table 5 where their private school attendance is indistinguishable from those who received a voucher in villages where private schools did not end up taking the voucher.

⁴⁶This is because in the random coefficients model lower asset families have less of a taste for private schools and, hence, less reason to search.

where w_i is a vector of weights and $L_{1i}(\theta)$ is i 's likelihood contribution given they are eligible (and $L_{0i}(\theta)$ is analogously defined). The weights account for sampling probabilities and attrition, as described elsewhere. For kindergartners, \tilde{e}_i is their observed AP voucher eligibility status; for first graders, \tilde{e}_i is their conditional or posterior probability of eligibility. This posterior eligibility probability is given by:

$$\tilde{e}_i = \frac{e_i L_{1i}(\theta)}{e_i L_{1i}(\theta) + (1 - e_i) L_{0i}(\theta)}$$

where e_i is the (logit) probability i is eligible. The algorithm iterates until the parameters converge.

We assume ineligible households already paid the search cost, so their likelihood contribution is equivalent to the control ability-to-pay constrained model. The likelihood contribution of eligible households reflects both ability-to-pay and search:

$$L_{i1}(\theta) = \sum_{j_i^*} \phi_{ij_i^*} \prod_{j \in \mathcal{V}_i} \left[\frac{1}{R} \sum_r P_{ij}^r(j_i^*) \right]^{d_{ij}}$$

where $\phi_{ij_i^*}$ is the probability the household can afford the schools in j_i^* . The numerical integration over the private school random effects is represented by the r superscript and the choice probability is given by:

$$P_{ij}^r(j_i^*) = \begin{cases} 0, & j \text{ private \& } p_j > p_{j_i^*+1} \\ P_{ij|S}^r(j_i^*) P_{iS}^r(j_i^*), & j \text{ private \& } p_j \leq p_{j_i^*+1} \\ P_{ij|\neg S}^r(j_i^*) (1 - P_{iS}^r(j_i^*)) + P_{ij|S}^r(j_i^*) P_{iS}^r(j_i^*), & j \text{ government} \end{cases}$$

$P_{ij|\neg S}^r(j_i^*)$ is the probability i chooses (government) school j if they do not search for private schools and $P_{iS}^r(j_i^*)$ is the probability that i searches given choice set j_i^* . The probability of searching is given by:

$$P_{iS}^r(j_i^*) = 1 - \exp[\pi_l + \pi_s \ln P_{iG|S}^r(j_i^*)]$$

where $P_{iG|S}^r(j_i^*)$ is the probability that i chooses *any* government school after searching. π_l and π_s are the location and scale, respectively, of the exponentially distributed search cost shock. Note that for treatment market kindergarten applicants (i.e. voucher winners and losers), this probability will embed their expectations that they will not have to pay tuition and fees at private schools and will receive an enrollment incentive. Ex-post, this expectation is not met for voucher losers.

6.3 Results and Fit

Table 7 presents selected parameter estimates of the unified model alongside those obtained from the control ability-to-pay model; the full set of indirect utility parameter estimates are presented in Table A11.⁴⁷ Households are estimated to be slightly more price sensitive in the unified model (row 1). Incorporating search also shrinks the private school valuation gap between voucher-ineligible and voucher-eligible students. With voucher-eligible students having to pay a search cost to learn about their match effects with private schools in the unified model, the corresponding valuation of private schools for these households increases.

Table 7: Estimates: Selected Parameters—Control Ability-to-Pay and Unified Models

	Control	Unified
<i>Indirect Utility</i>		
Tuition and fees (1000s of Rs.)	-1.28 (0.58)	-1.54 (0.07)
First stage residual	1.77 (0.63)	1.60 (0.07)
Enrollment incentive		2.26 (0.19)
Private school	11.35 (2.35)	9.74 (0.39)
× Eligible for AP voucher	-10.13 (1.76)	-5.51 (0.45)
Private random effect σ	2.66 (0.27)	1.77 (0.09)
<i>Search</i>		
Location		-0.24 (0.09)
Scale		0.36 (0.03)
<i>Ability-to-pay constraint</i>		
Intercept	2.96 (0.55)	3.39 (0.68)
Eligible for AP voucher	-1.29 (0.41)	-0.77 (0.36)
Asset factor	1.09 (0.23)	1.20 (0.28)
σ	1.34 (0.28)	1.48 (0.32)
N households	4,251	8,374
N observations	35,796	69,413

Notes: Table reports selected parameter estimates (and standard errors in parentheses) of control ability-to-pay constrained model (Control) and unified model estimated on entire dataset (Unified), including ability-to-pay constraint and search cost parameters. Parameter on total siblings in the constraint not reported; indirect utility parameter estimates for both models are reported in Table A11.

⁴⁷We use the baseline IVs for the results presented in the text; the results are virtually identical when using the “alternative” IV set that replaces the school-level cost proxy instrument with a cost index (see Table A12).

The coefficient on the enrollment incentive term of the unified model in Table 7 is large and positive. Recall that it is identified from the types of schools voucher winners attend relative to what we would expect based on the behavior of those in control villages. The fact that the magnitude of this coefficient is greater than the coefficient on tuition suggests that the enrollment incentive is capturing other differences beyond those implied by bargaining over the voucher surplus.

Table 8: Estimates: Ability-to-pay Constraint and Search Probability

	Share unable to pay for . . .				Search privates	
	<i>any</i> private		<i>priciest</i> private		Control	Unified
	Control	Unified	Control	Unified	Control	Unified
First graders						
Overall	0.09	0.05	0.18	0.10	1.00	0.53
Lower caste	0.13	0.07	0.25	0.11	1.00	0.43
Both parents completed primary	0.11	0.05	0.23	0.10	1.00	0.35
Asset level < 3	0.24	0.14	0.44	0.26	1.00	0.43
Asset level = 3	0.09	0.04	0.20	0.09	1.00	0.51
Asset level = 4	0.03	0.01	0.08	0.03	1.00	0.56
Asset level > 4	0.01	0.01	0.03	0.01	1.00	0.60
Voucher program applicants						
Control markets	0.13	0.06	0.25	0.13	1.00	0.47
Voucher losers	0.12	0.07	0.27	0.14	1.00	0.81
Voucher winners	0.16	0.08	0.33	0.17	1.00	0.82

Notes: Table reports estimates for shares of households constrained by ability-to-pay absent the voucher and who search private school options by subgroup per the estimates of the control ability-to-pay constrained model and unified model. Any and priciest private schools refer to among those in the household's village.

Table 8 reports the share of households that paid the search cost as well as the extent to which the ability-to-pay constraint is binding by subgroup. Voucher winners are substantially more likely to search for private schools than applicants in control villages, who are in turn more likely to search than eligible non-applicants in either control or treatment villages. Voucher losers in treatment villages are also more likely to search as we treat them as expecting to receive a voucher at the search stage in order to reconcile their high rates of private school attendance.

The estimates in Table 7 suggest that ability-to-pay is less constraining when search is accounted for and this is also reflected in the numbers in Table 8. The control model forces any search effects to instead operate through the ability-to-pay constraint. While less binding in general according the unified model, the constraint is still meaningful: 6-8% of voucher applicants cannot afford any private school and more than 13 percent cannot afford the most expensive private school in their village per the unified model estimates.

Given the estimates of the control and unified models, we next calculate willingness-to-pay for different school attributes and how this varies with household characteristics. Average marginal

willingness-to-pay for school characteristic k is given by:

$$\frac{1}{N} \sum_i \frac{\hat{\beta}_i^k}{-\hat{\alpha}}$$

where $\hat{\beta}_i^k$ is household i 's coefficient on characteristic k ; $\hat{\alpha}$ is the coefficient on tuition in the indirect utility function. Results are presented in Table A10. The unified model estimates that a one test score standard deviation increase in primary school math value-added has a present value of more than Rs. 800 to the average household (about 3% of annual median consumption per capita).⁴⁸ This value is nearly half as much as estimated by the control ability-to-pay constrained model. We also calculate willingness-to-pay for private schooling (and its associated change in school attributes) as:

$$\frac{1}{N} \sum_i \frac{\hat{\beta}_i(\bar{X}^P - \bar{X}^G)}{-\hat{\alpha}}$$

where \bar{X}^P and \bar{X}^G are the average characteristics of private and government schools, respectively. The table shows that willingness-to-pay for private schooling is similar between models, but the distributions across households is less extreme according to the unified model estimates as a result of accounting for search.

Table 9: Unified Model Goodness-of-Fit

	Attend Private Data	Private Unified	Tuition Private Data	Private Unified
First graders				
Overall	0.57	0.58	1.71	1.70
Lower caste	0.34	0.36	1.65	1.62
Both parents completed primary	0.27	0.28	1.48	1.60
Asset level < 3	0.28	0.33	1.45	1.57
Asset level = 3	0.52	0.54	1.72	1.70
Asset level = 4	0.68	0.66	1.84	1.76
Asset level > 4	0.78	0.78	1.67	1.69
Voucher program applicants				
Control markets	0.34	0.32	1.88	1.65
Voucher losers	0.45	0.45	2.08	1.90
Voucher winners	0.85	0.83	2.10	2.09

Notes: Table presents private school attendance and tuition given private school attendance by subgroup in the data with numbers implied by the unified model estimates to assess goodness-of-fit. Note that differences in the Data column with Table 1 are due to tabulating the numbers in this table using the sampling/attrition weights, which are used in the unified model estimation.

Table 9 shows that these two additional features—search costs and enrollment incentives—

⁴⁸To place in present value terms, we multiply the value reported in Table A10 by $\sum_{t=1}^5 \delta^{t-1}$, where $\delta = 0.90$ corresponds to an approximate 11% annual discount rate.

significantly improve the fit of the model, both with regard to the rate at which different groups attend private school but also by providing a better match with the posted tuition of the schools that voucher winners attend. The first set of columns shows that the predicted rates of private school attendance, both for voucher winners and voucher losers in treatment villages, are now within three percentage points of what is observed in the data. The second set of columns focuses on tuition conditional on attendance. Enrollment incentives are important here, with expected tuition now in line with what is observed for voucher winners who attend private schools.

6.4 Implications for Welfare

In this subsection, we use the model estimates to study welfare impacts of counterfactual voucher programs. To do so, we simulate the school choices of households in control villages given a voucher amount and eligibility criteria that we specify. We implement a voucher as a coupon provided to households that pays *up to* its full value towards tuition at any government-recognized private schools, thereby eliminating the enrollment incentives for low-tuition schools that were present in the original experiment. This exercise allows us to calculate effects on social welfare given the choice environments—household locations, school location, and school amenities—that exist in the data (i.e. the supply side is held fixed). This exercise thus abstracts away from strategic and general equilibrium adjustments that would be expected at scale.⁴⁹

We begin by considering a program that makes a voucher worth up to the 90th percentile of private school tuition for the duration of primary schooling universally available to all households. As Table 10 shows, our estimates predict that this program would raise the private school share from 58% to 75% of all households (17 points).

We then calculate complier households’ willingness-to-pay (WTP) for the program—given by the added inverse of the estimated compensating variation—by simulation. Letting d_{ij}^{r0} and d_{ij}^{r1} indicate choices absent the program (0) and under the program (1) for simulation r , we calculate:

$$\frac{1}{N_c} \frac{1}{R} \sum_i \sum_r \left[\frac{\sum_j d_{ij}^{r1} U_{ij}^{r1} - \sum_j d_{ij}^{r0} U_{ij}^{r0}}{-\alpha} \right] \cdot \mathbf{1} \left(\sum_j \text{PrivateVoucher}_j (d_{ij}^{r1} - d_{ij}^{r0}) > 0 \right)$$

which averages the scaled the expected utility changes among compliers by the coefficient on tuition;

⁴⁹By providing the voucher to households as a coupon, neither private schools nor households can keep any surplus of the voucher amount above the school’s price. For this reason, and because we would expect their effect to differ from the AP project in counterfactuals where siblings are also voucher-eligible, our simulations do not allow private schools to use voucher surplus to offer enrollment incentives.

R is the number of draws and N_c the number of compliers.⁵⁰ As Table 10 reports, we estimate that the average household who would otherwise attend a government school (i.e. the average complier) would be willing-to-pay Rs. 1,460 for the program, which in present value terms translates to about 25% of median annual per capita consumption.⁵¹ We also calculate and report how much program expenditure goes to defraying tuition paid by inframarginal households, i.e., those who would have chosen private schools even without the voucher:

$$\frac{1}{N_a} \frac{1}{R} \sum_i \sum_r [\sum_j d_{ij}^{r1} \max\{0, p_j - V\} - \sum_j d_{ij}^{r0} p_j] \cdot \mathbf{1}(\sum_j PrivateVoucher_j(d_{ij}^{r1} - d_{ij}^{r0}) = 0)$$

where V is the voucher amount (Rs. 2,600) and N_a is the number of always takers. Though the voucher funds complier households' choice to switch to a private school, we find that Rs. 1,690 on average are spent on every household who would have attended a private school anyway under a universal voucher.

Table 10: Welfare Impacts of Universal Voucher

Outcomes					
Increase in private schooling					0.17
Average complier HH's WTP (1000s of Rs.)					1.46
Average inframarginal HH's cost (1000s of Rs.)					1.69
Welfare metrics				Fiscal ext.	
				1/3	2/3
MVPF				1.33	3.05
Benefit/cost ratio				0.93	1.29
				0.70	1.06
				MEB	
				0.5	
				1.5	

Notes: Table reports welfare impacts of voucher program universally available to all households that covers tuition up to Rs. 2,600, as estimated by the unified model. 1000 simulations. Willingness-to-pay calculated by compensating variation. Two fiscal externality scenarios are considered: one where, for every household induced to switch to a private school (i.e. complier), 1/3rd of government spending per pupil (Rs. 8,400) is cut; another where 2/3rds is cut. Benefit/cost ratios calculated assuming marginal excess burdens (MEB) of 0.5 and 1.5.

In evaluating the overall welfare impact of the voucher program, the inefficiency of paying the tuition of inframarginal households must be weighed against the potential fiscal externality: how much of per pupil spending in government schools (about Rs. 8,390 in Andhra Pradesh per Dongre 2012) could be cut for every student who uses the voucher to exit government schooling? We consider two scenarios: a small impact scenario, where just 1/3rd could be recovered, and a large impact scenario, where 2/3rds—approximately equal to the share of spending allocated to

⁵⁰ $N_c = \sum_i \mathbf{1}(\sum_j PrivateVoucher_j(d_{ij}^{r1} - d_{ij}^{r0}) > 0)$

⁵¹We convert to present value by multiplying by $\sum_{t=1}^5 \delta^{t-1}$, where $\delta = 0.90$ corresponds to an approximate 11% annual discount rate.

teachers—could be recovered.

To summarize welfare impacts, we compute marginal values of public funds (MVPFs) as well as benefit/cost ratios. The calculations differ in their treatment of revenue and expenditure. For MVPFs, *net* spending enters the denominator:

$$\text{MVPF} = \frac{\overline{WTP}}{\overline{cost} - \text{Take-up} \times \psi \times 8.4}$$

When the denominator is negative, the program pays for itself and MVPF is defined as ∞ . \overline{WTP} and \overline{cost} are the *average* household’s willingness-to-pay and tuition paid for by the program. We find MVPFs of 1.33 and 3.05 for $\psi = 1/3$ and $\psi = 2/3$, respectively (Table 10). For benefit/cost ratios, revenue enters the numerator, and spending (which must be appropriated) the denominator:

$$\text{Benefit/cost ratio} = \frac{\overline{WTP} + \text{Take-up} \times \psi \times 8.4 \times (1 + MEB)}{\overline{cost} \times (1 + MEB)}$$

We consider two values for marginal excess burden (MEB) of taxation, the economic cost to society of raising tax revenue. An MEB of 0.5 implies that each dollar of revenue imposes an additional efficiency cost of 50 cents. An MEB of 1.5, by contrast, is highly distortionary, indicating that the welfare cost exceeds the revenue raised. If the fiscal externality is large, we find that the benefit/cost ratio of the universal voucher program exceeds one in either case, implying that the program improves social welfare.

How do these welfare estimates compare with those from the control models? Appendix Table A13 presents comparisons, showing that only the unified model estimated on both control and treatment data finds that a universal voucher increases social welfare (given a large fiscal externality and large marginal excess burden). How it differs from the respective control model estimates is revealing. The control ability-to-pay constrained model arrives at large welfare metrics (e.g. MVPF of 2.14) but for the wrong reasons: it ascribes even greater WTP to complier households, who its estimates are more constrained than the unified model, while predicting only somewhat less take-up (1-2 points less). In contrast, the control random coefficient model predicts significantly less take-up (7 points lower) and estimates the average complier’s WTP is over 25% lower.

While the unified model results highlight scenarios in which a universal voucher program raises welfare, it stands to reason that targeted programs could be more efficient by minimizing voucher use by inframarginal households. With this motivation, we consider a counterfactual program offering a voucher of the same maximum value (Rs. 2,600), but which is only available to the

bottom quarter of households in terms of socioeconomic status (i.e. own two or fewer assets). Table 11 reports welfare impacts from this program.⁵² Though its impact on private schooling is smaller (7 points), the average complier’s willingness-to-pay is comparable with the universal case. However, far less of each inframarginal household’s tuition gets paid for by the targeted program. The result is greater efficiency. The MVPF is nearly three in the small fiscal impact scenario and equals infinity—the targeted voucher program is self-financing—when the fiscal externality is large. In this latter scenario, the targeted program would be expected to generate up to \$2 in social welfare for every \$1 in tuition the program pays for.

Table 11: Welfare Impacts of Voucher Targeted to Asset-Poor Households

Outcomes						
Increase in private schooling						
0.07						
Average complier HH’s WTP (1000s of Rs.)						
1.52						
Average inframarginal HH’s cost (1000s of Rs.)						
0.22						
Welfare metrics				Fiscal ext.		
				1/3	2/3	
MVPF				2.93	∞	
Benefit/cost ratio						
				MEB	0.5	1.27
				1.5	1.05	1.76

Notes: Table reports welfare impacts of voucher program targeted only to households with two or fewer assets that covers tuition up to Rs. 2,600, as estimated by the unified model. 1000 simulations. Willingness-to-pay calculated by compensating variation. Two fiscal externality scenarios are considered: one where, for every household induced to switch to a private school (i.e. complier), 1/3rd of government spending per pupil (Rs. 8,400) is cut; another where 2/3rds is cut. Benefit/cost ratios calculated assuming marginal excess burdens (MEB) of 0.5 and 1.5.

Lastly, we examine welfare impacts if the voucher could be reliably targeted to ability-to-pay constrained households, according to the unified model estimates. To do so, we consider a voucher amount to each household, \tilde{V}_i , parameterized as:

$$\tilde{V}_i = \min\{\kappa \times \hat{\phi}_{i1} \times 8.4, V\}$$

where $\hat{\phi}_{i1}$ is the posterior probability that household i cannot choose *any* private school in their village. 8.4 represents government spending per pupil. κ thus governs the slope of the voucher amount with respect to the ability-to-pay constraint and the max voucher amount remains as before, Rs. 2,600. We consider voucher programs that vary κ from 0.5 up to a value of 4. The results in Table 12 show that targeting ability-to-pay directly is appreciably more efficient than the coarse targeting of asset levels. In particular, a program that sets $\kappa = 3$ yields more take-up while

⁵²Note that our simulations do not consider the effect targeted programs may have on the tuition charged by private schools to non-eligible students.

also achieving a MVPF that is twice as large. Flatter slopes on the voucher policy generate even greater efficiency gains (to a smaller set of households) yet. These results highlight the empirical importance of ability-to-pay constraints in this setting and their influence on how much welfare a voucher program can create.

Table 12: Welfare Impacts of Vouchers Targeted to Ability-to-Pay Constrained Households

	$\kappa =$	0.5	1	2	3
Increase in private schooling		0.03	0.05	0.07	0.08
MVPF (Fiscal ext. = 1/3)		∞	23.17	7.66	6.12

Notes: Table reports welfare impacts of voucher programs targeted based on ability-to-pay if implemented in control markets, as estimated by the unified model. 1000 simulations. The voucher formula is $\tilde{V}_i = \min\{\kappa \times \hat{\phi}_{i1} \times 8.4, V\}$, where $V = 2.6$. MVPFs assume that for every household induced to switch to a private school (i.e. complier), 1/3rd of government spending per pupil (Rs. 8,400) is cut; MVPF = ∞ for all κ if 2/3rds is cut.

Our counterfactuals take the sets and attributes of schools as given. But as we have seen through the enrollment incentives, how a voucher program is designed will invariably lead to supply-side responses. Vouchers paid directly to schools at a fixed high amount created incentives for private schools to recruit enrollees, distorting household choices. On the other hand, tying the voucher to actual tuition of the school, as under India’s Right to Education (RtE) Act, may lead to responses both on quality and price of existing schools, as in Sahai (2025), as well as potentially inducing entry.⁵³ Such supply side-responses may improve outcomes for those who take-up vouchers but, due to rising prices, may be welfare decreasing for ineligible private school students.

A further consideration for targeting scaled-up voucher policies is to ensure that disadvantaged households can access the voucher. In our setting, the project team visited eligible households in person to assist with applications. However, the RtE Act in practice has been shown to be regressive because asset poor and less educated households find it much more difficult to complete the paperwork needed to apply for the program (Romero and Singh, 2024).

7 Conclusion

Our paper makes both substantive and methodological contributions. Substantively, we show that a model of school choice with ability-to-pay constraints, search costs, and supply-side responses matches the high voucher take-up rates observed in the AP School Choice project. We estimate

⁵³The combined roles of entry and pricing decisions in shaping outcomes under voucher policies are studied by Sánchez (2023) in the context of Chile.

substantial welfare gains from counterfactual voucher programs in part due to the costs of government schools being significantly higher than their private counterparts. Further, our results show that the gain in consumer surplus is economically meaningful for many students induced into private schools by vouchers because of the presence of ability-to-pay constraints that otherwise constrain some households from consuming school quality up to the value of the numeraire good.

Methodologically, our research design isolates the contribution of experimental data to model development, identification, and selection. Out-of-sample model validation was a limited success: while the control models successfully fit the out-of-sample choice patterns of “control” households in treatment markets, all of the models underpredict experimental take-up of the voucher offer. In developing the models, we anticipated that this would most likely reflect inadequate instruments (or a mis-specified control function). Instead, data from treatment markets revealed other issues as first-order. Blinded to treatment data patterns, we failed to anticipate the intervention’s effects on household decision-making and school responses. Moreover, our later empirical quantification of these mechanisms relies on variation across both treatment and control markets for identification. Our experience suggests that when interventions alter the structure of household or firm behavior—which, as our case illustrates, can arise even in interventions that appear limited in scale—credible policy analysis will likely require estimating models that incorporate those effects using treatment or policy variation.

References

- Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters.** 2020. “Do parents value school effectiveness?” *American Economic Review*, 110(5): 1502–1539.
- Agte, Patrick, Claudia Allende, Adam Kapor, Christopher Neilson, and Fernando Ochoa.** 2024. “Search and biased beliefs in education markets.” National Bureau of Economic Research.
- Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja.** 2017. “Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets.” *The American Economic Review*, 107(6): 1535–1563. Publisher: American Economic Association.
- Angrist, Joshua, Eric Bettinger, and Michael Kremer.** 2006. “Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia.” *The American Economic Review*, 96(3): 847–862.
- Arcidiacono, Peter, Karthik Muralidharan, Eun-young Shim, and John D Singleton.** 2021. “Experimentally validating welfare evaluation of school vouchers: Part i.” National Bureau of Economic Research.
- Arteaga, Felipe, Adam J Kapor, Christopher A Neilson, and Seth D Zimmerman.** 2022. “Smart matching platforms and heterogeneous beliefs in centralized school choice.” *The Quarterly Journal of Economics*, 137(3): 1791–1848.
- ASER.** 2018. *Annual Status of Education Report*. Aser Centre.
- Attanasio, Orazio P., Costas Meghir, and Ana Santiago.** 2012. “Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA.” *The Review of Economic Studies*, 79(1): 37–66.
- Attansio, Orazio, and Debbie Blair.** 2018. “Structural Modelling in Policymaking.”
- Barseghyan, Levon, Maura Coughlin, Francesca Molinari, and Joshua C. Teitelbaum.** 2021. “Heterogenous Choice Sets and Preferences.”
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan.** 2007. “A Unified Framework for Measuring Preferences for Schools and Neighborhoods.” *Journal of Political Economy*, 115(4): 588–638.
- Ben-Akiva, Moshe, and Bruno Boccara.** 1995. “Discrete choice models with latent choice sets.” *International Journal of Research in Marketing*, 12(1): 9–24.
- Berry, Steven, James Levinsohn, and Ariel Pakes.** 1995. “Automobile Prices in Market Equilibrium.” *Econometrica*, 63(4): 841–890.
- Berry, Steven T, and Philip A Haile.** 2024. “Nonparametric identification of differentiated products demand using micro data.” *Econometrica*, 92(4): 1135–1162.
- Cameron, Stephen V., and James J. Heckman.** 2001. “The Dynamics of Educational Attainment for Black, Hispanic, and White Males.” *Journal of Political Economy*, 109(3): 455–499. Publisher: The University of Chicago Press.

- Carneiro, Pedro, Jishnu Das, and Hugo Reis.** 2022. “The value of private schools: Evidence from Pakistan.” *Review of Economics and Statistics*, 1–45.
- Chetty, Raj, John N. Friedman, and Jonah E. Rockoff.** 2014. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates.” *The American Economic Review*, 104(9): 2593–2632.
- Corradini, Viola.** 2023. “Information and access in school choice systems: Evidence from New York City.” Working paper.
- Currie, Janet, and Firouz Gahvari.** 2008. “Transfers in Cash and In-Kind: Theory Meets the Data.” *Journal of Economic Literature*, 46(2): 333–383.
- Delavande, Adeline, and Basit Zafar.** 2019. “University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints.” *Journal of Political Economy*, 127(5): 2343–2393.
- Deming, David J.** 2014. “Using School Choice Lotteries to Test Measures of School Effectiveness.” *American Economic Review*, 104(5): 406–411.
- Dongre, A.** 2012. “What is the Per Child Expenditure in Government Schools?” In *Accountability Initiative*. New Delhi.
- Duflo, Esther, Rema Hanna, and Stephen P. Ryan.** 2012. “Incentives Work: Getting Teachers to Come to School.” *The American Economic Review*, 102(4): 1241–1278.
- Epple, Dennis, Richard E. Romano, and Miguel Urquiola.** 2017. “School Vouchers: A Survey of the Economics Literature.” *Journal of Economic Literature*, 55(2): 441–492.
- Ferreira, Maria Marta.** 2007. “Estimating the effects of private school vouchers in multidistrict economies.” *American Economic Review*, 97(3): 789–817.
- Filmer, Deon, and Lant H. Pritchett.** 2001. “Estimating Wealth Effects Without Expenditure Data—Or Tears: An Application To Educational Enrollments In States Of India.” *Demography*, 38(1): 115–132.
- Galiani, Sebastian, Alvin Murphy, and Juan Pantano.** 2015. “Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment.” *The American Economic Review*, 105(11): 3385–3415.
- Galiani, Sebastian, and Juan Pantano.** 2023. “Structural Models.” In *Handbook of Labor, Human Resources and Population Economics*. 1–55. Springer.
- Gregory, Jesse.** 2017. “The Impact of Post-Katrina Rebuilding Grants on the Resettlement Choices of New Orleans Homeowners.”
- Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger.** 2005. “Parental Preferences and School Competition: Evidence from a Public School Choice Program.” National Bureau of Economic Research Working Paper 11805.
- Hausman, Jerry A.** 1996. “Valuation of New Goods under Perfect and Imperfect Competition.” In *The Economics of New Goods.*, ed. Timothy F. Bresnahan and Robert J. Gordon, 207–248. University of Chicago Press.

- Kamat, Vishal, and Samuel Norris.** 2025. “Estimating welfare effects in a nonparametric choice model: The case of school vouchers.” *arXiv preprint arXiv:2002.00103*.
- Kane, Thomas J., and Douglas O. Staiger.** 2008. “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation.” National Bureau of Economic Research Working Paper 14607.
- Keane, Michael P., and Kenneth I. Wolpin.** 2001. “The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment.” *International Economic Review*, 42(4): 1051–1103. Publisher: [Economics Department of the University of Pennsylvania, Wiley, Institute of Social and Economic Research, Osaka University].
- Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff.** 2015. “Value-added modeling: A review.” *Economics of Education Review*, 47: 180–195.
- Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E Waugh.** 2023. “The welfare effects of encouraging rural–urban migration.” *Econometrica*, 91(3): 803–837.
- Larroucau, Tomás, Ignacio Rios, Anaïs Fabre, and Christopher Neilson.** 2024. “College Application Mistakes and the Design of Information Policies at Scale.”
- Manski, Charles F., and Steven R. Lerman.** 1977. “The Estimation of Choice Probabilities from Choice Based Samples.” *Econometrica*, 45(8): 1977–1988. Publisher: [Wiley, Econometric Society].
- McFadden, Daniel.** 1978. “Modelling the Choice of Residential Location.” In *Spatial Interaction and Planning Models*, ed. F. Snickars and J. Weibull, 823–848. North-Holland, Amsterdam.
- 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.
- Neal, Derek, and Joseph Root.** 2024. “The Provision of Information and Incentives in School Assignment Mechanisms.” National Bureau of Economic Research.
- Neilson, Christopher.** 2013. “Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students.”
- Nevo, Aviv.** 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” *Econometrica*, 69(2): 307–342.
- Pathak, Parag A, and Peng Shi.** 2021. “How well do structural demand models work? Counterfactual predictions in school choice.” *Journal of Econometrics*, 222(1): 161–195.
- Petrin, Amil.** 2002. “Quantifying the Benefits of New Products: The Case of the Minivan.” *Journal of Political Economy*, 110(4): 705–729.
- Petrin, Amil, and Kenneth Train.** 2010. “A Control Function Approach to Endogeneity in Consumer Choice Models.” *Journal of Marketing Research*, 47(1): 3–13.
- Romero, Mauricio, and Abhijeet Singh.** 2024. “The Incidence of Affirmative Action: Evidence from Quotas in Private Schools in India.”
- Sahai, Harshil.** 2025. “School Voucher Design and Strategic Pricing: Evidence from India.”

- Schorfheide, Frank, and Kenneth I. Wolpin.** 2012. “On the Use of Holdout Samples for Model Selection.” *American Economic Review*, 102(3): 477–481.
- Schorfheide, Frank, and Kenneth I. Wolpin.** 2016. “To hold out or not to hold out.” *Research in Economics*, 70(2): 332–345.
- Sánchez, Cristián.** 2023. “Equilibrium Consequences of Vouchers Under Simultaneous Extensive and Intensive Margins Competition.”
- Tarozzi, Alessandro, Aprajit Mahajan, Brian Blackburn, Dan Kopf, Lakshmi Krishnan, and Joanne Yoong.** 2014. “Micro-loans, Insecticide-Treated Bednets, and Malaria: Evidence from a Randomized Controlled Trial in Orissa, India.” *American Economic Review*, 104(7): 1909–1941.
- 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.” *The American Economic Review*, 96(5): 1384–1417.
- Todd, Petra E, and Kenneth I Wolpin.** 2023. “The best of both worlds: combining randomized controlled trials with structural modeling.” *Journal of Economic Literature*, 61(1): 41–85.

A Appendix

A School Value-Added

We estimate each schools' value-added to student learning using baseline and follow-up exam scores in math and control for it in the choice estimation. We assume that the achievement of student i in subject k at year t is a linear function of household i 's characteristics, H_i , the quality of the school they attend ω_j , and their prior exam performance:

$$A_{ijt} = \rho(A_{it-1}) + H_i' \pi + \omega_j + \zeta_{ijt} \quad (11)$$

A_{ijt} is the student's exam performance in year t , which is normalized across students within year. School j 's unobserved value-added to the learning process, ω_j , is assumed to be fixed within our panel. We include a cubic of prior exam performance, A_{ijt-1} , and control for student demographics, parental education, and household socioeconomic status in H_i (e.g. age, whether female, whether scheduled caste, whether both parents completed primary education, asset levels, etc.).

We observe up to three math scores for each student, with the first two drawn from *non-consecutive* school years t (the baseline year and year three) in the data. We therefore estimate school value-added by first estimating equation (11) separately by wave w :

$$A_{ijw} = \rho_w(A_{iw-1}) + H_i' \pi_w + \omega_{jw} + \zeta_{ijw}$$

For $w = 3$, A_{iw-1} is i 's test score from the prior school year, while A_{iw-1} for $w = 2$ corresponds to i 's score at baseline. This estimation step thus yields up to two fixed effects for each school, $\hat{\omega}_{j2}$ and $\hat{\omega}_{j3}$. We then shrink the fixed effect estimates using empirical Bayes techniques to reduce the impact of measurement error (Kane and Staiger, 2008; Deming, 2014; Koedel, Mihaly and Rockoff, 2015).

To finally recover the time-invariant value-added ω_j for each school, we combine the wave 3 posterior with an assumption that wave 2 is inflated by some factor $\delta > 1$ that is common to all schools and households. Specifically, when both measurements are available, we average $\tilde{\omega}_{j2}/\delta$ and $\tilde{\omega}_{j3}$.⁵⁴ We estimate δ by iterating on two steps: First, given δ^m , averaging the value-added estimates with a fixed effect regression:

$$\begin{pmatrix} \tilde{\omega}_{j2}/\delta^m \\ \tilde{\omega}_{j3} \end{pmatrix} = \omega_j + \lambda_w + \epsilon_{jw}$$

Second, running $\tilde{\omega}_{j2}/\delta^m = \beta\hat{\omega}_j + \epsilon_2$ and updating δ^m to $\delta^{m+1} = \hat{\beta}\delta^m$, which enforces the condition that the wave 2 estimates are the same *on average* as each school's underlying value-added (i.e. at convergence, $\hat{\beta} = 1$). We find that this is satisfied for $\hat{\delta} = 1.56$

⁵⁴Note that if $\rho()$ is linear, then $\delta = 1 + \rho$.

B Weights

We weight students’ likelihood contributions in estimation to account for three aspects of the samples’ construction. First, students who did not apply for the AP project voucher (kindergartners and first graders, the latter asked regarding a voucher they might receive for the next year, i.e. second grade) are underrepresented in the samples. We adjust for this using the student counts reported in Table II of Muralidharan and Sundararaman (2015) to estimate sampling probabilities.

Second, first graders (whose primary school choices were collected at baseline) were sampled conditional on their choice of primary school. For consistency, we therefore re-weight to match the population market shares of students attending public and private schools (Manski and Lerman, 1977). To do so, we consult India’s Annual Status of Education Report (ASER) survey in 2008 (the same calendar year at the project baseline). We calculate population shares of private school attendance at the district level, restricting the ASER sample to households in Andhra Pradesh villages with at least one private school and students in the age range of our estimation sample (and excluding children not enrolled in school).

Finally, there is attrition of kindergartners from baseline.⁵⁵ We regard as attritors students who do not attend a primary school in their village. If no choice is available in the first wave of household tracking data for control data kindergartners, we use the next wave, when available, to record choices; this applies to 483 of the 1,485 students in our control markets kindergartner subsample. We adjust for attrition with inverse probability weights, estimated using a probit model. The attrition probability depends on student and household sociodemographic characteristics, baseline Telugu score, district of residence, as well as whether they are eligible and/or applied for the AP project voucher. For estimation of the unified model, we analogously construct weights for the non-voucher winner kindergartners in treatment markets. Estimation of the unified model also applies weights to the voucher winning kindergartners to adjust for the voucher offer lowering attrition (discussed in Appendix D).

C Coding Voucher Take-Up

To code voucher take-up in a way consistent with how the counterfactual simulations were done using the control markets, we combine information from tracking and from the project team. Table A6 summarizes the coding. Our starting place is the majority of students (416) labeled as accepting the offer and who attend a private school in tracking data. We add to this group 69 students who attend a private school in the tracking data. As Table A6 shows, most of these are students who later “dropped out” of the voucher program or who were deemed ex-post ineligible due to gaining admission to a private school prior to learning their voucher outcome. We further code as voucher users 21 students who were unable to use the voucher by virtue of being too young (irrespective of where tracking data show them attending school).⁵⁶ For eventual dropouts, we assume that the

⁵⁵Because first graders were sampled at their primary school choice, attrition is 0 by construction.

⁵⁶There is also one student with the extenuating circumstance of “waiting list not used” that we code as intending to use.

private school observed in the tracking data is their initial voucher school. For the cases where the tracking data does not align with our coding of use, we otherwise keep our analyses agnostic about precisely where students would’ve used the voucher.

Importantly, in eight treatment market villages, no students randomized-in to receive an offer actually used a voucher due to non-compliance by private schools.⁵⁷ For purposes of the experimental validation, we remove these non-complying villages from the sample entirely.⁵⁸ This leaves a sample of 574 households who were randomly offered a voucher in treatment villages. Of these, 489 (85%) intended to take-up the voucher offer. Later on, we look at choice patterns of households in the non-compliant treatment villages to test mechanisms that could explain underprediction of take-up.

D Voucher Offers Lowered Attrition

The (kindergartner) attrition rate, calculated as the share of households at baseline who attend a primary school in their village, is noticeably smaller for households offered a voucher (11%) than it is for control market applicants (19%). The voucher offer appears to have increased household retention in the final sample by attracting students to local private schools. We thus adjust model predictions and estimates based on the treatment data to account for selective attrition. To do this, we first solve for the number of households that would have attrited from the final offered student sample *in the absence of the voucher offer*. The calculation assumes that the attrition rates of applicants between treatment and control markets would be the same in the absence of the offer and comes to 70 of the 574 students (in non-flagged villages). We then assume that the 70 students who otherwise would have attrited also belong to the subgroup of students who actually took-up the voucher offer.

Under this assumption, we use the calculation of excess attriters in two ways in the analysis. First, we handicap control model predictions for voucher take-up for selective attrition by adding 70 students to the number of voucher users predicted by the control models. Second, we assign weights to the students who actually used the voucher such that they effectively represent 70 fewer students. We also adjust the weights to account for differences in the probability of attrition between those students (as a function of observed characteristics). These weights are applied when using offered households’ choice data to test control model fit and in estimation of the unified model.

⁵⁷This can be seen in Table A6, where excluding these “flagged” villages removes all of the offered students coded as “school rejected” from the sample. Several private schools in otherwise compliant treatment villages also reneged on participating. This was in part because of the project requirement that all students in these schools should take independent learning assessments (this concern was raised by other private schools too, which is why the assessments used in Muralidharan and Sundararaman (2015) were conducted outside school). We therefore do not set tuition and fees to zero at these specific schools when generating model predictions.

⁵⁸We also flag and remove one additional treatment village where take-up was not zero but was abnormally low.

E Specification Tests of Control Ability-to-Pay Constrained Model

Table A9 presents results of specification tests of the ability-to-pay constrained model pre-committed to in Arcidiacono et al. (2021). Column (1) shows that with no controls the model significantly underestimates intended usage, while column (2) shows that a large and significant voucher-eligible private school fixed effect rationalizes the underprediction. Column (4) shows that the underestimate of take-up is especially pronounced at lower tuition schools. Columns (5) and (6) indicate that the underprediction is also pronounced for those with older siblings at government schools. Column (7) of Table A9 adds interactions between students’ baseline Telugu scores and school characteristics—a private school intercept, whether English medium, estimated math value-added, and whether offers Hindi instruction—to the model to test for ability sorting. The control models did not include ability heterogeneity, largely due to these measures being unavailable for first graders. The results suggest that some mis-specification may come from greater take-up among higher ability students, but higher ability students actually “prefer” lower value-added schools (and vice versa). Column (8) allows for the possibility that offered students value voucher-eligible private schools’ attributes differently than implied by the control models. These results indicate higher disutility of travel to voucher schools, much weaker preferences for English medium instruction and for value-added, and greater preferences for Hindi classes. While interesting, the inclusion of these covariates does little to explain the overall underprediction of voucher take-up nor does their inclusion meaningfully modify the negative coefficient on voucher school tuition.

F Unified Model Instruments and Sample

We estimate the first stage of private schools’ tuition and fees on observed school characteristics and instruments on the full sample of private schools from both control and treatment markets. The first stage does not allow for heterogeneity by village treatment status, consistent with an assumption that the intervention did not impact tuition-setting. The estimates are shown in Table A4. The results in the text use the baseline IVs (column 3); results using that substitute the cost proxy with the (quadratic) cost index are very similar. The estimates imply that the average treatment village private school is unobservably better than the average control private school.

The estimation sample for the unified model combines all of the subgroups shown in Figure 1. Households in non-complying treatment villages are included in the estimation sample, but voucher winners in these non-compliant villages are treated like voucher losers in the estimation. Likewise, for those households we code as using the voucher but who did not actually use one to attend a private school in the data, their observed (non-voucher) school choice is matched with the one predicted by being a household that expects a voucher but ex-post does not receive one.

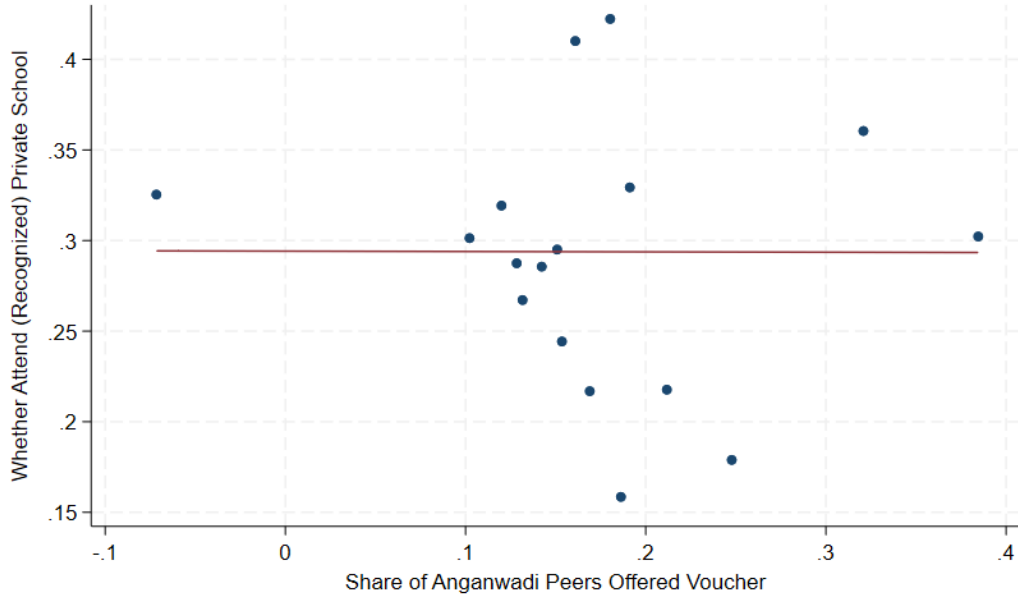
G Additional Figures and Tables

Table A1: Summary Statistics: Household Characteristics by Subgroup

	First Graders				Applicants		Kindergartners Non-applicants		Ineligible	
	Attend Gov't Mean	Diff	Attend Private Mean	Diff	Mean	Diff	Mean	Diff	Mean	Diff
Female	0.52	0.02	0.47	0.02	0.58	-0.02	0.55	0.07	0.47	-0.00
Lower caste	0.34	0.01	0.12	-0.01	0.32	0.03	0.36	-0.02	0.11	-0.02
Muslim	0.06	-0.00	0.09	-0.01	0.07	0.02	0.07	-0.06*	0.08	0.02
Christian	0.07	0.01	0.04	-0.01	0.08	0.01	0.11	-0.02	0.04	0.02*
# siblings	2.37	0.01	2.18	-0.12**	2.23	0.05	2.29	-0.08	2.13	-0.03
Older sibling in gov't school	0.50	0.01	0.11	-0.06***	0.37	-0.00	0.48	0.02	0.10	-0.03
Both parents completed primary	0.09	-0.00	0.34	-0.03	0.17	0.01	0.15	-0.02	0.35	-0.01
≥ 1 parent completed secondary	0.06	0.00	0.25	-0.04	0.10	0.00	0.07	-0.01	0.25	-0.05
Both parents laborers	0.45	-0.01	0.18	0.04*	0.39	0.00	0.43	-0.05	0.19	-0.03
Math score z (baseline)	0.02	0.01	0.64	0.14**						
Telugu score z (baseline)	0.03	0.07**	0.72	-0.03	0.00	0.04	-0.04	-0.42***	0.39	-0.15**
Owns home	0.75	0.01	0.76	0.05*	0.76	-0.01	0.76	-0.00	0.77	0.00
Pucca house	0.72	0.01	0.92	-0.02	0.75	0.01	0.65	0.03	0.91	-0.00
Water facility in home	0.41	-0.01	0.60	-0.04	0.44	-0.07***	0.45	-0.05	0.61	-0.08**
Household toilet	0.24	-0.02	0.58	-0.00	0.28	-0.03	0.23	0.04	0.57	0.05
Owns land	0.18	0.02**	0.31	-0.02	0.19	-0.01	0.17	0.09*	0.33	0.02
Asset level < 3	0.39	-0.02	0.13	0.02	0.36	0.04	0.40	-0.06	0.12	0.01
Asset level = 3	0.27	0.00	0.21	-0.02	0.26	-0.02	0.26	-0.01	0.20	-0.03
Asset level = 4	0.20	0.02	0.29	-0.03	0.23	-0.01	0.23	0.04	0.27	0.00
Asset level > 4	0.13	0.00	0.37	0.02	0.15	-0.01	0.11	0.04	0.40	0.01
First principal asset factor	-0.13	0.01	0.43	-0.06	-0.05	-0.04	-0.15	0.06	0.44	-0.01
N households	4439		975		1915		258		787	

Notes: Table reports summaries of household characteristics by subgroups as well as treatment-control balance checks. Means and N refer to all households in control and treatment markets; columns labeled “Diff” report differences in means (and their statistical significance) between households in the subgroup in treatment markets and in control markets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Figure A1: Testing for Peer Effects of Voucher Offers on Non-Offered Applicants



Notes: Figure presents residualized binscatter of whether a non-offered applicant kindergartner student attends a government-recognized private school against the share of peers in their Anganwadi who were offered a voucher as a test for peer effects. Sample includes control applicants—all of whom had 0 peers offered a voucher by constructions—and applicant not offered a voucher in treatment markets (who do not reside in a non-complying village). Regression controls for whether the village is a treatment market or not, the share of Anganwadi peers who applied for the voucher, and a quadratic in the number of Anganwadi peers.

Table A2: Summary Statistics: Characteristics of Primary Schools

	Government		Private	
	Mean	Diff	Private	Diff
Tuition and fees (Rs.)	0.81	-1.45	1924	226**
English medium	0.02	0.00	0.57	-0.08*
Unrecognized	0	.	0.23	-0.04
Mid-day meals	0.99	0.00	0.03	-0.01
Kitchen facility	0.26	0.04	0.01	-0.00
Full pucca building	0.89	-0.01	0.52	0.08**
Library	0.94	-0.01	0.77	-0.01
Functional water tap	0.42	0.05	0.62	0.02
Functioning toilet	0.65	0.01	0.84	0.05
Separate toilet for girls	0.34	0.07*	0.60	0.02
Staffroom for teachers	0.20	0.00	0.72	0.03
Playground	0.52	0.00	0.70	0.04
Has secondary school	0	.	0.27	0.05
Total school enrollment	74.28	-1.88	286.18	8.69
Multi-class teaching	0.70	0.10***	0.24	-0.06*
Pupil-teacher ratio	26.53	1.00	16.68	1.20
Share teachers absent	0.21	-0.04***	0.09	-0.01
Share teachers with BA	0.78	-0.00	0.54	-0.05
Share teachers with formal certificate	0.90	0.01	0.16	0.01
Share teachers female	0.50	-0.07***	0.71	-0.01
Share teachers lower caste	0.24	-0.02	0.12	0.01
Share teachers Muslim	0.02	-0.01	0.07	-0.01
Share teachers from village	0.25	0.03	0.48	0.02
Offers Hindi instruction	0	.	0.44	0.02
Offers computer skills	0.01	0.01	0.13	-0.00
School value-added	-0.04	0.02	0.04	-0.05
N schools		686		570

Notes: Table reports summaries of school characteristics by government and private as well as treatment-control balance checks. Means and N refers to all schools; columns labeled “Diff” report differences in means (and their statistical significance) between schools in treatment markets and in control markets. ** p<0.01, * p<0.05, . p<0.1

Table A3: Preference Heterogeneity in Baseline Model Specifications

Control Models		Unified	School characteristic	Interactions
RC	CC			
X			Tuition and fees	AP elig., total siblings, discrete asset levels (e.g. asset level = 3, etc.)
X	X	X	Distance	AP elig., female, Muslim, lower caste
			Closest Public	
X	X	X	Private	AP elig., female, Muslim, lower caste, parental education, older sibling gov't, total siblings
X	X	X	English medium	Female, Muslim, lower caste, parental education
X	X	X	Value-added	Female, Muslim, lower caste, parental education
X	X	X	Offers Hindi	Female, Muslim, lower caste, parental education
			Facilities factor	
			Teaching quality factor	
			Teacher Char. factor	
			Has secondary school	

Table details preference heterogeneity in baseline model specifications discussed in the text. Parental education variables are whether both parents completed primary and whether at least one completed secondary school. Facilities (e.g. pucca), teaching quality (e.g. multiclass teaching, share teachers absent), and teacher characteristics (e.g. share female, from village) attributes are first principal factors. Variables included in the models but not listed in the table are indicators for imputed tuition and fees and for missing value-added, effects of which are not heterogeneous across households.

Table A4: First Stage: Private School Tuition and fees

	(1) Control Markets		(3) Control + Treatment	
	Baseline IVs	Alternative	Baseline IVs	Alternative
Product space location	238.1*** (56.42)	288.9*** (59.60)	197.0*** (44.07)	210.1*** (44.45)
Cost proxy	0.376*** (0.135)		0.155* (0.080)	
Cost index		0.246*** (0.0947)		0.418*** (0.1105)
Cost index ²		-0.000599*** (0.000132)		-0.000717*** (0.000138)
First-stage F	12.51	17.63	11.39	20.64
Cragg-Donald stat	11.20	13.13	11.39	16.29
R ²	0.309	0.341	0.232	0.265
Observations	293		570	

Notes: Table presents first stage estimates that regress private school tuition and fees on school characteristics and instrumental variables on the control markets sample (columns 1 and 2) and the entire sample (columns 3 and 4). Baseline IVs refers to instruments summarizing product space location (first factor of fixed characteristics of other private schools in *same* village) and proxying for school-level costs (predicted tuition and fees based on similar private schools in *other* villages), while Alternative replaces the cost proxy with a village-level cost index (and its square). Estimation and validation results of control models in the text pertain to column (1); unified model estimation uses column (3). Though not reported, regressions control for the school characteristics included in the choice models. Standard errors reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A5: Estimates: Selected Parameters—Control Models

	RC	CC
Tuition and fees (1000s of Rs.)	-2.35 (0.28)	-1.28 (0.58)
× Eligible for AP voucher	0.07 (0.12)	
× Asset level = 2	0.45 (0.20)	
× Asset level = 3	0.74 (0.20)	
× Asset level = 4	1.12 (0.20)	
× Asset level > 4	0.81 (0.21)	
First stage residual	1.60 (0.20)	1.77 (0.63)
Private random effect σ	2.23 (0.22)	2.66 (0.27)
<i>Ability-to-pay constraint</i>		
Intercept		2.96 (0.55)
Eligible for AP voucher		-1.29 (0.41)
Asset factor		1.09 (0.23)
σ		1.34 (0.28)

Notes: Table reports selected parameter estimates (and standard errors in parentheses) of control random coefficient model (RC) and control ability-to-pay constrained model (CC). Coefficient on total siblings in ability-to-pay constraint excluded from the table. The estimation sample contains 4,251 households and 35,796 household-school observations. All indirect utility estimates for both models are reported in Arcidiacono et al. (2021).

Table A6: Coding Voucher Take-up

Voucher code	Tracking	N	N*	Use
accepted and admitted	Private	416	410	yes
	Government	9	9	no
rejected voucher	Private	8	8	yes
	Government	49	49	no
migrated	Private	1	1	yes
	Government	9	9	no
own private admission	Private	31	22	yes
	Government	12	11	no
under age	Private	7	6	yes
	Government	14	14	yes
admitted, dropped out	Private	29	27	yes
	Government	7	7	no
waiting list not used	Private	0	0	.
	Government	1	1	yes
school rejected	Private	9	0	.
	Government	27	0	.
Total		629	574	489

Notes: Table displays coding of voucher take-up based on information from project team (Voucher code) and tracking data. N represents counts of households in each cell; N* reports counts excluding households residing in nine “flagged” treatment villages where, collectively, very few students were actually able to use a voucher to attend a private school.

Table A7: Validation: Voucher Elasticity of Private Schooling

	RCT	RC	CC
Overall	221	116	148
Female	252	126	159
Muslim	110	72	85
Lower caste	328	158	209
Older sibling in gov’t school	474	262	335
Both parents completed primary school	116	87	110
≥ 1 parent completed secondary	66	63	78
Both parents laborers	259	132	176
Asset level < 3	303	189	247
Asset level = 3	190	125	162
Asset level = 4	247	87	124
Asset level > 4	136	78	78

Notes: Table presents average voucher elasticity (percent change in private schooling due to the voucher offer) of applicant households by subgroup in the treatment data (RCT), and as predicted by the random coefficient (RC) and ability-to-pay constrained control models (CC). Predictions correspond to baseline specification described in the text and detailed in Arcidiacono et al. (2021).

Table A8: Validation: Voucher Intent-to-Treat Effects and Elasticities on Characteristics of Chosen School

	RCT		RC		CC	
	ITT	ϵ	ITT	ϵ	ITT	ϵ
Tuition and fees (Rs.)	1.08***	183	0.68	120	0.94	168
English medium	0.13***	54	0.08	42	0.14	72
Distance to school (mi.)	-0.25	-21	-0.15	-15	-0.15	-14
School value-added	0.01		0.00		0.01	
Offers Hindi	0.33***	206	0.11	59	0.17	90
Unobservable	0.25***		0.08		0.07	

Notes: Table presents voucher intent-to-treat effects (ITT) and elasticities (ϵ) – the percent change in the average value of the choice characteristic due to the voucher offer – for applicant households in the treatment data (RCT), and as predicted by the random coefficient (RC) and ability-to-pay constrained control models (CC). Predictions correspond to baseline specifications described in the text and detailed in Arcidiacono et al. (2021). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A9: Validation: Hypothesis Tests for Mis-specification of Ability-to-Pay Constrained Control Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private school		2.60		5.28	4.53	4.70	4.58	3.98
		(0.22)		(0.40)	(0.42)	(0.50)	(0.43)	(0.48)
Private school \times Asset factor					0.03	-0.76		
					(0.31)	(0.63)		
Private school \times Older sibling in gov't school					1.63	0.89	1.72	1.74
					(0.42)	(0.81)	(0.43)	(0.45)
Tuition and fees (1000s of Rs.)			0.52	-1.32	-1.34	-1.43	-1.37	-1.26
			(0.08)	(0.16)	(0.16)	(0.21)	(0.16)	(0.18)
Tuition and fees \times Asset factor						0.39		
						(0.26)		
Tuition and fees \times Older sibling in gov't school						0.36		
						(0.33)		
Private voucher school \times Telugu z (baseline)							0.47	0.27
							(0.26)	(0.29)
English medium \times Telugu z (baseline)							-0.37	-0.34
							(0.28)	(0.29)
Value-added \times Telugu z (baseline)							-1.08	-1.08
							(0.29)	(0.31)
Offers Hindi \times Telugu z (baseline)							0.00	-0.12
							(0.32)	(0.34)
Distance \times Private school								-0.70
								(0.19)
English medium \times Private school								-0.66
								(0.32)
Value-added \times Private school								-2.09
								(0.63)
Has Hindi \times Private school								0.63
								(0.37)
First stage residual \times Private school								-0.08
								(0.21)
\hat{U}_{se}	0.65	0.84	0.74	0.84	0.84	0.84	0.84	0.84
AIC	1,400	1,235	1,360	1,164	1,153	1,153	1,137	1,105

Notes: Table reports hypothesis tests of model mis-specification that examine variables' predictive power for voucher winners' choice patterns conditional on the indirect utility of the alternative implied by the control ability-to-pay constrained model estimates. Note that both the private school indicator and tuition and fees are zeroed out for voucher-ineligible private schools (i.e. those without government recognition) for these analyses. Standard errors reported in parentheses.

Table A10: Estimates: Willingness-to-Pay (1000s of Rs.) for School Characteristics—Control Ability-to-Pay and Unified Models

	English Medium		School Value-Added		Private Schooling	
	Control	Unified	Control	Unified	Control	Unified
Overall	0.83	0.73	0.37	0.20	3.20	3.69
Female	0.43	0.50	0.39	0.21	2.89	3.65
Muslim	1.90	1.41	0.28	0.47	4.33	3.90
Lower caste	0.68	0.60	0.37	0.00	0.37	2.42
Older sibling in gov't school	0.65	0.62	0.37	0.19	-0.36	1.67
Both parents completed primary school	1.72	1.37	0.40	0.27	6.45	5.52
≥ 1 parent completed secondary	1.98	1.62	0.18	0.13	7.19	5.65
Both parents laborers	0.49	0.50	0.39	0.15	0.88	2.58
Asset level < 3	0.57	0.56	0.38	0.17	0.62	2.45
Asset level = 3	0.74	0.67	0.38	0.20	2.52	3.40
Asset level = 4	0.92	0.78	0.37	0.21	3.99	4.05
Asset level > 4	1.06	0.89	0.35	0.22	5.39	4.72

Notes: Table reports point average willingness-to-pay (1000s of Rs.) by subgroup for English medium instruction, a one test score standard deviation increase in school value-added, and private schooling as estimated by the control ability-to-pay constrained (Control) and unified (Unified) models. Willingness-to-pay for private schooling calculated as change from average government to average private school. Estimates correspond to baseline specifications described in the text. Control numbers differ from Arcidiacono et al. (2021) because no adjustment to place in present value terms is made here.

Table A11: Estimates: Indirect Utility Parameters—Control Ability-to-Pay and Unified Models

	Control		Unified	
	Coef	SE	Coef	SE
Tuition and fees (1000s of Rs.)	-1.28	0.58	-1.54	0.07
First stage residual	1.77	0.63	1.60	0.07
Private random effect σ	2.66	0.27	1.77	0.09
Enrollment incentive			2.26	0.19
Log distance	-1.41	0.09	-0.69	0.06
× Eligible for AP voucher	0.29	0.15	-0.58	0.06
× Age > 5	0.15	0.08	0.02	0.05
× Female	-0.14	0.08	-0.04	0.05
× Muslim	0.13	0.14	0.19	0.08
× Lower caste	-0.05	0.09	0.03	0.05
Private school	11.35	2.35	9.74	0.39
× Eligible for AP voucher	-10.13	1.76	-5.51	0.45
× Female	-0.60	0.25	-0.20	0.13
× Muslim	0.16	0.46	-0.01	0.23
× Lower caste	-1.50	0.29	-0.74	0.14
× Both parents completed primary	0.17	0.42	0.76	0.20
× ≥ 1 parent completed secondary	0.58	0.53	0.10	0.25
× Older sibling in gov't	-2.59	0.49	-2.15	0.12
× Total siblings-2	-0.07	0.10	-0.17	0.07
English medium	0.90	0.40	0.95	0.10
× Female	-0.92	0.23	-0.64	0.11
× Muslim	1.42	0.47	1.07	0.18
× Lower caste	0.05	0.27	-0.03	0.14
× Both parents completed primary	0.85	0.29	0.68	0.15
× ≥ 1 parent completed secondary	1.35	0.60	1.31	0.17
Unrecognized private school	-0.61	0.16	-1.03	0.08
Value-added	0.51	0.17	0.33	0.20
× Female	0.03	0.19	0.05	0.10
× Muslim	-0.16	0.29	0.37	0.19
× Lower caste	-0.04	0.20	-0.37	0.05
× Both parents completed primary	0.31	0.25	0.28	0.17
× ≥ 1 parent completed secondary	-0.51	0.28	-0.32	0.19
Offers Hindi	0.03	0.33	-0.09	0.11
× Female	0.55	0.34	0.10	0.13
× Muslim	1.15	0.43	0.63	0.21
× Lower caste	1.15	0.38	0.16	0.16
× Both parents completed primary	0.39	0.30	0.13	0.17
× ≥ 1 parent completed secondary	0.18	0.33	0.09	0.19
Closest public school	0.78	0.12	0.59	0.06
Facilities factor	0.46	0.07	0.32	0.03
Teaching quality factor	-0.34	0.05	-0.14	0.04
Teacher characteristics factor	-0.06	0.04	-0.08	0.02
N households	4,251		8,374	
N observations	35,796		69,413	

Notes: Table reports point estimates (and standard errors) for indirect utility parameters of control ability-to-pay constrained model (Control) and unified model estimated on full dataset (Unified). Estimates on indicator for whether school serves secondary grades, whether value-added is missing, whether tuition is imputed, and whether distance is missing not reported.

Table A12: Estimates: Selected Parameters—Unified Model: Baseline and Alternative IVs

	Baseline IVs	Alternative
<i>Indirect Utility</i>		
Tuition and fees (1000s of Rs.)	-1.54 (0.07)	-1.48 (0.10)
First stage residual	1.60 (0.07)	1.52 (0.10)
Enrollment incentive	2.26 (0.19)	2.18 (0.20)
Private school	9.74 (0.39)	9.50 (0.60)
× Eligible for AP voucher	-5.51 (0.45)	-5.25 (0.58)
Private random effect σ	1.77 (0.09)	1.52 (0.21)
<i>Search</i>		
Location	-0.24 (0.09)	-0.21 (0.08)
Scale	0.36 (0.03)	0.35 (0.03)
<i>Ability-to-pay constraint</i>		
Intercept	3.39 (0.68)	3.53 (0.79)
Eligible for AP voucher	-0.77 (0.36)	-0.63 (0.38)
Asset factor	1.20 (0.28)	1.30 (0.33)
σ	1.48 (0.32)	1.57 (0.38)

Notes: Table reports selected parameter estimates (and standard errors in parentheses) of unified model using baseline IVs and alternative IVs. Baseline IVs refers to instruments summarizing product space location (first factor of fixed characteristics of other private schools in *same* village) and proxying for school-level costs (predicted tuition and fees based on similar private schools in *other* villages), while Alternative replaces the cost proxy with a village-level cost index (and its square). Sample contains 8,374 households and 69,413 household-school observations. Parameter on total siblings in the constraint not reported.

Table A13: Comparing Universal Voucher Welfare Estimates Between Control Models and Unified Model

	Control		
	RC	CC	Unified
Outcomes			
Increase in private schooling	0.10	0.16	0.17
Average complier HH's CV	1.06	2.05	1.46
Average inframarginal HH's cost	1.81	1.85	1.69
Welfare metrics			
MVPPF	1.28	2.14	3.05
Benefit/cost ratio	0.73	0.95	1.06

Notes: Table reports welfare impacts of voucher program universally available to all households that covers tuition up to Rs. 2,600 if implemented in control markets, as estimated by the control random coefficient (RC) and ability-to-pay constrained (CC) models and the unified model (Unified). 1000 simulations. Willingness-to-pay calculated by compensating variation. Welfare metrics calculations assume that, for every household induced to switch to a private school (i.e. complier), 2/3rds of government spending per pupil (8,400 Rs.) is cut; benefit/cost calculated assuming marginal excess burden (MEB) of 1.5.