The incidence of affirmative action: Evidence from quotas in private schools in India*

Mauricio Romero† Abhijeet Singh‡

June 27, 2023

Abstract

The incidence of redistributive policies is central to whether they meet their stated goals. We study this in the context of one of the world's largest programs to improve social equity in schooling: a 25% quota in all Indian private schools for students from disadvantaged groups. We use lottery-based estimates to show that, although students admitted under the quota attend more expensive and preferred schools on average, the distribution of program benefits is very regressive. Program applicants are concentrated among more-educated and better-off households. Consequently, 7.4% of the program spending accrues to the bottom socioeconomic quintile, compared to 24.3% to the top quintile. In total, two-thirds of the per-child cost of a quota seat is inframarginal for school choice. We use rich survey data to show that low application rates for poorer children are not driven by preferences and beliefs. Instead, information constraints and application frictions appear to be key. Finally, we use a randomized intervention to confirm the importance of these frictions and further demonstrate that alleviating a single constraint (e.g., information) may not reduce regressive selection, even if it boosts application rates substantially. Our results demonstrate how constraints facing potential applicants can make redistributive policies regressive in practice. Appropriate policy interventions must consider the joint incidence of these constraints to reduce regressivity.

JEL Codes: I24, I25, I28, O15, H42

*This project was executed in partnership with Indus Action and the Government of Chhattisgarh. Ramamurthy Sripada, Dwithiya Raghavan, Archana Prabhakar, and Tanisha Agrawal provided excellent field management and research assistance. We are grateful to Orazio Attanasio, Maria Atuesta, Antonella Bandiera, Natalie Bau, Peter Bergman, Jishnu Das, Jean Dreze, Tore Ellingsen, Esther Gehrke, Peter Hull, Mariana Laverde, Clare Leaver, Erik Lindqvist, Craig McIntosh, Karthik Muralidharan, Christopher Neilson, Robert Östling, Lant Pritchett, Gaurav Khanna, Imran Rasul, Evan Rose, Santiago Saavedra, Miguel Urquiola, Diego Vera-Cossio and seminar participants at various institutions for comments and suggestions. Financial support from the RISE Programme, funded by the UK's Foreign, Commonwealth and Development Office, is gratefully acknowledged. Romero gratefully acknowledges financial support from the Asociación Mexicana de Cultura, A.C. All errors are our own. We received IRB approval from the Institute for Financial Management and Research. The RCT in this paper was registered as AEARCT-0010428.

†ITAM; J-PAL; E-mail: mtromero@itam.mx

‡Stockholm School of Economics; J-PAL; E-mail: abhijeet.singh@hhs.se

1 Introduction

Governments routinely attempt to promote social equity by prioritizing poorer citizens in the delivery of essential services such as housing, healthcare and education. Where private providers account for a large market share, common interventions to equalize access include targeted subsidies, price caps, or legal requirements to include otherwise under-served groups. In practice, however, these policies may disproportionately benefit better-off households, undermining stated goals of fairer allocation.[1] Further, since taxpayers prioritize fairness and redistribution concerns when evaluating policy proposals (Stantcheva, 2021), such regressivity could also undermine popular support. Thus, a central concern for all social policy is to assess the *de facto* incidence of program benefits and, in many instances, how to make it more progressive.

We study these issues in the context of one of the world's largest affirmative action policies in schooling: quotas in private schools in India. In India, as in other low- and middle-income countries, fee-charging private schools are common and differentiated in price, quality and amenities (Muralidharan and Sundararaman, 2015; Urquiola, 2016; Andrabi et al., 2022). The Right to Education (RTE) Act 2009 imposed a quota of 25% of the student intake in all Indian private schools for students from disadvantaged economic and caste backgrounds to reduce socioeconomic stratification. The state pays tuition fees for children enrolled under this quota (up to a cap) and schools are not allowed to select which students they admit or charge top-up fees. In 2018/19, approximately 4 million students nationwide used a quota seat to attend private schools (Indus Action, 2019).

We use data from Chhattisgarh, a state of $\sim$29 million people that has implemented the quota policy since 2010, to ask three related questions. First, we evaluate the effect of receiving a quota seat on students' schooling choices (the primary goal of the policy). We identify this effect using lotteries embedded in the mechanism used by the state to allocate quota seats. Second, we evaluate the extent to which the policy is redistributive in practice by estimating the proportion of applications and expenditure that accrue to different socioeconomic groups; we do this by comparing the pool of all policy applicants to state-wide representative data using a comparable measure of socioeconomic status. Finally, we shed light on the factors underlying differential take-up of the policy across the socioeconomic distribution. We do this by using: i) detailed survey data on parental preferences, beliefs, and knowledge of the policy, ii) a randomized intervention to relax application constraints, and iii) administrative data to understand spatial constraints related to the availability of private schools.

[1] For example, rent control in housing markets is often (ostensibly) motivated to protect poorer tenants and reduce socioeconomic stratification but may end up primarily benefiting more-privileged households (Glaeser and Luttmer, 2003; Diamond et al., 2019; Ahern and Giacoletti, 2022).

We analyze the effects of winning quota seats on schooling choices using data on all eligible quota applications in 2019, supplemented with survey data on educational choices and socioeconomic background. Eligible students can apply for admission under the quota in either preschool grades (Nursery and Kindergarten) or in Grade 1, the formal start of primary schooling, depending on their age and whether the private school offers preschool education or not.[2] Seats are allocated using a centralized algorithm (Immediate Acceptance) and we identify program effects using the lottery-based allocation of oversubscribed slots (Abdulkadiroğlu et al., 2017).

Our first result focuses on the effects of winning a quota seat on the extensive margin of private school enrollment. Being allotted an RTE quota seat increases the probability of attending any private school by 25 percentage points (p-value < 0.001). This is over a base of $\sim$75% of applicants who were *not* assigned an RTE seat but attend private schools by paying regular fees (i.e., virtually all students offered an RTE seat attend private schools). Much of the extensive margin effect is concentrated in preschool grades, covered by the quota in schools with integrated preschool sections, shifting students from home care to preschool. By Grade 1, when enrollment is compulsory and near-universal, this effect is 13 percentage points (p-value < 0.001).

Next, we focus on shifts within the private sector to schools with different characteristics (i.e., the intensive margin of private school choice). Many applicants would eventually attend the same private school with or without a quota seat: In Grade 1, $\sim$50% of applicants who lose the lottery for their top choice nevertheless attend the same school as fee-paying students. Consequently, a quota seat does not lead lottery winners to attend schools with a significantly different caste composition of the student body. However, some students use the quota seat to upgrade within the private sector. Schools attended by lottery winners are more likely to have English-medium instruction, rank higher in parents' applications, and are more expensive. Students with an RTE seat in Grade 1 attend schools that, on average, charge 48% higher annual fees (a USD 38 increase over a control mean of USD 79). This treatment effect is 50% of the average fee reimbursed by the state — thus, half of the public spending per seat is effectively a cash transfer to households. The full cost per seat is higher since 30% of schools charge fees higher than the government reimbursement cap (of $\sim$93 USD per year). Valuing quota seats at the full price paid by non-quota students, about 67% of the expenditure on a quota seat is inframarginal for school choice, the primary targeted outcome for the policy.[3]

[2]Applications for non-quota private school seats are decentralized. Students may seek admission for a non-quota seat in any private school and, if admitted, pay the regular tuition fees to attend.

[3]The expenditure is completely inframarginal for students who would have attended the same school without an RTE seat. For other treated students, the inframarginal portion equals the difference between fees reimbursed by the state and what they would have paid (potentially in a different school) without the policy.

This high degree of inframarginality is surprising for a policy targeted at poorer disadvantaged groups, given prior evidence of steep socioeconomic gradients in private school enrollment (e.g., Bagde et al. (2022)). A possible explanation is substantial selection, within targeted groups, into the pool of applicants for quota seats. We investigate this by comparing measures of socioeconomic status of all applicants (beyond those applying to oversubscribed schools) with identical measures in state-wide representative data. Quota applicants have more household assets and parental education than the average household with similarly-aged children in the state, both within eligible caste groups and in the overall population. We compute differential rates of take-up and program expenditure across the socioeconomic distribution: 7.4% of the program spending accrues to the bottom socioeconomic quintile, while 24.3% of spending accrues to households in the top quintile.

Reflecting the centrality of regressive selection into applying for the program in determining the incidence of program benefits, the rest of the paper focuses on understanding why application rates are particularly low for poorer households. This could inform the design of potential interventions to improve progressivity.[4] We mainly focus on three demand-side explanations (Currie, 2006): (a) that poorer parents have low demand for private schools; (b) that they lack information about the policy; and (c) that they face greater administrative burdens and application complexity.[5]

We measure parental demand and information by collecting rich survey data from randomly-sampled households with a child aged between 3–7 years in urban and rural areas of Raipur, the most populous district in Chhattisgarh. We elicit parental demand for private schools, and their expectations of student experience, using stated choice exercises similar to Delavande and Zafar (2019).[6] Demand for private schools is high even in the bottom quintile and well above application rates for the RTE quota. Parents across the socioeconomic distribution expect their children to have better experiences and outcomes in private schools. Further, in line with the policy's motivation, financial constraints deter a substantial fraction of below-median households from accessing private schools. Thus, low parental demand for private schools does not seem to rationalize the low application rates of poor households.

[4]This exercise could also explain why a large share of seats go unfilled nationally. In 2019–20, only 56% of seats available in Chhattisgarh were filled, a rate comparable to other major states (Indus Action, 2019).

[5]In Section 5.2, we further investigate a supply-side explanation: that poor neighborhoods are less likely to be served by private schools. While such spatial segregation is important, we show that substantial reallocation is still possible.

[6]Specifically, we elicit rankings over fictitious private and public schools in scenarios that vary the out-of-pocket fees of different private schools. We also elicit expectations of children's experiences in these schools. See Section 4 for details and Appendix C for exact questions and the validation of the survey data. This exercise follows work using directly-elicited measures of preferences and expectations to study education choices in high school and college (Arcidiacono et al., 2012; Attanasio and Kaufmann, 2014; Boneva and Rauh, 2017; Delavande and Zafar, 2019).

In contrast, information constraints and barriers to applying are much more salient. Only 20% of households at the bottom-end of the socioeconomic distribution report having heard of the policy, compared to 65% at the top. Poorer households are also less likely to have internet access and are less familiar with navigating online application portals. We use a field experiment to investigate whether relaxing these constraints would suffice to address the regressive selection of applicants. We randomly selected 459 households in our household sample, out of 914 households with mobile phones, to receive detailed information about the policy (such as application deadlines, eligibility criteria, and required documentation), combined with in-person assistance for submitting online applications. The intervention was delivered in a two-week window while applications were open.

The intervention boosted application rates for the RTE quota by 9.5 percentage points (p-value .0037), a substantial increase of 43% over a control mean of 22%. However, it did not reduce regressive selection: although information constraints and barriers to applying online were more severe for poorer households, treatment gains were not larger. This counter-intuitive result reflects a further regressive constraint: the administrative burden of certifying eligibility. Richer households were more likely to have all documents — 56% for the top decile compared to 24% in the bottom decile — which largely reflects difficulty in obtaining documents rather than "true" ineligibility. Thus, poorer households were disproportionately less able to apply, despite being willing to, and the intervention did not increase the proportion of poorer households in the applicant pool.

These experimental results, combined with the baseline survey evidence, highlight the sources of regressive selection and the likely (in)effectiveness of many approaches to resolving this. Low-income eligible parents are willing to apply for quota seats but are constrained by information about the policy and the requirements of the application process. The inability of willing poor households to submit applications provides a stark illustration of how an "O-ring" process (Kremer, 1993) contributes to regressive selection. Successful applications require *all* constraints to be satisfied simultaneously, and all constraints in our setting affect poorer households more. Consequently, untargeted interventions that do not solve all frictions simultaneously would fail to address regressive selection. Common interventions, such as chatbots and information helplines, may even reduce the share of the poorer households in the applicant pool despite successfully boosting overall application rates (as we find in our experiment). We discuss the potential for alternative intervention designs to improve progressivity in Section 5.

Our principal contribution is to provide evidence on one of the world's most ambitious policies to promote social mobility through school education: 4 million students attend private elementary schools using RTE quota seats, with projections of up to 16 million if the program was fully implemented in all states (Indus Action, 2019). Through potential effects on peers (see Rao (2019) for evidence of effects on non-beneficiaries of this policy), this could affect the entirety of the private schooling sector, which accounted for ∼95 million children in K-12 education in 2020-21.[7] Despite the importance of this policy, however, empirical evidence of its effects on direct quota beneficiaries has remained scarce.[8] We provide evidence on the current effects of the policy on school choice, on the effectiveness of public spending and the incidence of its benefits, and on potential supplementary interventions that could boost its redistributive potential.

This evidence speaks to global debates and relates closely to three strands of research on promoting social equity in education. First, we contribute to the literature evaluating interventions to promote school integration (see, e.g., Bergman (2018); Campos and Kearns (2022)) and a broader literature studying school segregation and stratification.[9] Our findings are directly relevant in highlighting constraints affecting current school integration initiatives (and measures to improve them even within current policy rules). Improvements may benefit both poor and rich students in our setting: private schools in India have shown positive effects on skills with large labor market returns (Muralidharan and Sundararaman, 2015; Singh, 2015), and interaction with poorer students generates positive peer effects in the behavior of richer students (Rao, 2019).

Second, we relate to an extensive literature on affirmative action in education (e.g., Arcidiacono and Lovenheim (2016); Bagde et al. (2016); Bleemer (2022); Otero et al. (2021); Mello (2021)). In contrast to the exclusive focus on college admissions in this work, we provide unique evidence at *school entry age*. This distinction is important because concerns of academic mismatch and controversy about the fairness of "non-merit" admissions criteria are both more muted at the start of formal education than in college.

Third, our results also relate directly to evidence on other efforts to equalize access to quality education by using public funds for private provision. For example, vouchers and private management of public schools are common throughout the world (e.g., Epple et al. (2017); Romero et al. (2020); Cohodes and Parham (2021)). RTE quotas

[7] As a benchmark for scale, the ∼4 million students enrolled on RTE quotas in 2019 exceed total enrollment in all Charter Schools in the US (∼3 million, (Cohodes and Parham, 2021)) or any voucher program internationally (Epple et al., 2017). Enrollment in Indian private schools is larger than in most education systems.

[8] The voucher experiment in Muralidharan and Sundararaman (2015) was inspired by the (not-then-enacted) proposals for the RTE quotas. However, as a stand-alone intervention with substantial outreach through household visits, it did not study incidence or effects on school choice in steady-state (our primary focus). See also Damera (2017) who studies RTE quotas in Karnataka but does not study the incidence of public spending, or investigate reasons for undersubscription.

[9] See, for instance, a vast literature on racial and socioeconomic segregation in US schools (e.g., Clotfelter (2011); Reardon and Owens (2014)) and in other settings (e.g., the UK (Jenkins et al., 2008), Chile (Hsieh and Urquiola, 2006), and Scandinavia (Söderström and Uusitalo, 2010)).

represent an alternative mechanism in this class of policies and we show that regressive take-up blunts the redistributive goals of the policy. Regressive selection may make these policies less progressive in practice than intended, albeit for different reasons. For instance, Walters (2018) documents that disadvantaged students have low demand for Charter Schools, despite having larger gains from attending. In our setting, the joint incidence of multiple constraints yields an outcome that is observationally equivalent (undersubscription by poorer households) but low demand is not the principal reason that deters take-up. Thus, understanding the causes of regressive take-up is central to identifying appropriate interventions to boost progressivity.

Finally, going beyond education, we also directly contribute to the literature studying the take-up of social programs across sectors. Our results contrast with influential studies in development economics where self-targeting and ordeals improve progressivity, or where low adoption rates are caused by low demand (Mobarak et al., 2012; Cole et al., 2013); instead, they resemble results from studies of application frictions in social programs in the US.[10] Results from our randomized intervention, highlighting multiple constraints, are perhaps closest to Banerjee et al. (2021) who find similar results with interventions to broaden enrollment in universal health insurance programs in Indonesia. However, we show that the incidence of multiple constraints affects not just overall application rates but also the *composition* within the eligible groups (which is central to redistributive programs and, here, also leads directly to inframarginality of program spending). Thus, optimal interventions to address regressive selection will likely differ from those focused only on raising program take-up, a distinction that is critical but underemphasized. These insights are likely to also generalize to redistributive policies in other domains including, for example, quotas in higher education and employment (Bagde et al., 2016; Khanna, 2020).

2 Policy Background and Context

2.1 Private schooling in India

Fee-charging private schools in India account for ∼31% of primary school enrollment in rural areas and ∼50% in urban areas (Pratham, 2019). At the primary level, private schools typically have larger enrollment and more teachers than government schools. They are also more likely to use English as a medium of instruction and offer single-grade classrooms and subject-specific teachers. Further, private primary schools often include an integrated preschool section. Teachers in private schools are

[10]For US-based evidence on the importance of application frictions, see e.g., Currie (2006); Finkelstein and Notowidigdo (2019); Deshpande and Li (2019); Bhargava and Manoli (2015). For evidence from developing countries on self-targeting and ordeals, see e.g., Ravallion (1991); Besley and Coate (1992); Alatas et al. (2016). The framework in Finkelstein and Notowidigdo (2019) can reconcile these contrasting results.

typically younger, less likely to have completed formal teacher training, and are paid substantially less than teachers in government schools. However, they are less likely to be absent from school or engaged in non-teaching activities than government school teachers. These broad characterizations are consistent across multiple national-level datasets (see Kingdon (2020) for an informative review).

Evidence on relative differences in quality is more scarce. Cross-sectionally, students in private schools score higher than those in private schools but this difference is confounded by selection: private school students are more likely to be from higher socioeconomic backgrounds, urban, and male — they are also likely selected on various unobserved characteristics. Muralidharan and Sundararaman (2015) provide the best estimates of the causal effects of attending a private school in rural Andhra Pradesh state and find that there are large effects on English and the national language (Hindi) with no evidence of lower effects in math and the local language; similar results are shown by Singh (2015) using value-added models in the same state. Unfortunately, similar evidence, whether using experimental variation or individual-level panel data, is unavailable in other states. There is no causal evidence that looks at the distribution of effects within the private sector.[11]

2.2 Quotas in the Right to Education Act 2009

The Right to Education (RTE) Act, enacted in 2009 by the national Parliament, sets the regulatory framework for organizing the entire school system, and made free and compulsory education from 6–14 years a fundamental right.

We focus on Clause 12(1)(c) of the act, which established the 25% quota in private schools. This provision was motivated by concerns that the rapid growth of fee-charging private schools led to segregated schools and classrooms, and impeded access to high-quality schooling for students from disadvantaged backgrounds. Guidelines for implementing the act stress "the need for moving towards composite classrooms with children from diverse backgrounds, rather than homogeneous and exclusivist schools", which echoes school integration reforms elsewhere. The clause requires fee-charging private schools to "admit at least 25% of the strength of class I, children belonging to weaker section and children belonging to disadvantaged group from the neighborhood and provide them free and compulsory education till completion of elementary education. Further, where the school admits children at pre-primary level, such admissions will be made at that level".[12]

[11] Indeed, the only such evidence from a similar setting is from three districts in rural Punjab in Pakistan by Andrabi et al. (2022). They find substantial heterogeneity in the productivity of schools within the public and the private sector. Private schools have higher value-added, on average, but there is substantial overlap across sectors.

[12] "Weaker section" in the law typically refers to income-poor households, and "disadvantaged groups" to castes and tribal groups that have historically been discriminated against. The quota additionally covers

The government reimburses private schools for tuition fees and other expenditures on students admitted through this quota at notified levels.

The use of quotas to broaden access reflects a long history of affirmative action programs in India: it resembles, for example, caste-specific quotas in political representation (Pande, 2003; Jensenius, 2015; Bhavnani, 2017), college admissions (Bertrand et al., 2010; Bagde et al., 2016) and government employment (Prakash, 2020) or gender-specific quotas in village councils (Beaman et al., 2012; Chattopadhyay and Duflo, 2004). The program resembles voucher and charter school systems in aims and the reimbursement of school fees by the state. Unlike vouchers, the benefits provided by a quota seat are school-specific and not transportable to other schools. Unlike Charter schools, schools are only reimbursed for (up to) a quarter of their enrollment and lottery-based enrollment procedures also only apply to a subset of the enrolled class.

This provision has been contentious and was litigated up to the Supreme Court, which affirmed its constitutionality in 2012. However, as with many desegregation policies elsewhere — for example, *Brown vs. Board of Education* in the US — universal adoption did not immediately follow the ruling. Individual states in India retain substantial *de facto* power to decide in whether (and how) to implement the quotas, such as defining admissions procedures, the rules for reimbursement and the precise definition of eligible groups. Thus, adoption has been partial and staggered across states; the policy remains unimplemented in many states. In 2018–19, ∼4 million students were enrolled in an RTE quota seat; full national implementation would cover an estimated ∼16 million children annually (Indus Action, 2019).

2.3 Quotas in Chhattisgarh: context and lottery design

Our study is based in Chhattisgarh state, which had a population of ∼29.4 million in 2020 and has historically been disadvantaged across several development indicators. In 2011, ∼40% of the population was estimated to be below the poverty line (compared to ∼22% nationally). In 2019, the national government ranked the state 21 (out of 28 states) in its achievements of the UN's Sustainable Development Goals (NITI Aayog, 2020).

In Chhattisgarh, in 2012, ∼24% of overall primary school enrollment was in private schools but, as in the rest of the country, there was substantial variation across socioeconomic groups. About 79% of children aged 7-12 in the top decile of the SES distribution (measured by consumption per capita) were enrolled in private schools, compared to 4% in the bottom decile (see Figure A.1a). For children from Scheduled

children with medically-attested physical and/or mental disabilities, orphaned children, and HIV-positive children. These latter groups account for a very small share of applicants.

Castes and Scheduled Tribes, which are historically disadvantaged caste groups, these figures were 18% and 7% respectively (see Figure A.1b).[13]

Chhattisgarh has implemented the RTE-mandated quota since 2010. Children are eligible for an RTE seat if they are aged 3–7, and meet one of two criteria: (i) either their household is classified as "economically weaker" based on official documentation;[14] Or, (ii), the households belong to Scheduled Castes (SC) or Scheduled Tribes (ST).[15] The government reimburses school fees for students admitted under the quota up to a cap of INR 7,000, and provides student grants for books and uniforms. Schools cannot charge top-up fees (even if the school fees exceed the cap for reimbursements).

Applications for quota seats were decentralized (i.e., submitted to schools individually) from 2010 to 2017 when the state introduced a centralized online application portal.[16] The allocation process remained decentralized in 2017 and 2018 (even if the applications were not) and moved to state-level centralized allocation from 2019 onwards (the year for which we use the data). 40% of applicants used caste as the basis for eligibility ("Disadvantaged Groups") and the rest used income ("Economically Weaker Sections") (Indus Action, 2021).

The state used an Immediate Acceptance algorithm to allocate quota seats in 2019. Specifically, the algorithm is as follows:

1. Parents rank as many private schools in their catchment area as they want, in their order of preference. They may apply for their child's admission to one of three grades — Nursery, Kindergarten (KG) and Grade 1 — as determined by their child's age. Eligible ages are 3–4 years for Nursery, 4–5 years for KG, and 5–6.5 years for Grade 1.
2. All students are assigned to their first-preference school if it is not over-subscribed. No priority is given to students with enrolled siblings, living nearby, or otherwise.
3. Students whose first-preference school is over-subscribed enter a lottery (separate for each grade). Each child is in only one school-grade lottery per round.
4. Schools with filled quotas and allocated students are removed.

[13] We use the Indian Human Development Survey (IHDS)-II (2011-12) to compute figures for Chhattisgarh because it includes both urban and rural areas and has information on socioeconomic characteristics, caste, and school type (if enrolled). See https://doi.org/10.3886/ICPSR36151.v6 for more details.

[14] In Chhattisgarh, this requires having access to a card certifying "Below Poverty Line" status as determined in administrative surveys in 2002 (rural) and 2007 (urban); or the Socio-Economic Caste Census in 2011; or to have a ration card for the Antyodaya Anna Yojana (AAY) which is given to particularly poor households.

[15] These groups are recognized by the Constitution of India as historically disadvantaged. They are entitled to affirmative action in political representation, education, and employment. 43.4% of the population in Chhattisgarh in 2011 belonged to these groups compared to about one-quarter nationally. Access to caste-based government benefits requires official documentation of caste ("caste certificate") issued by local officials.

[16] Applicants without access to internet or computers can also apply in person at specified government offices with assistance. These applications are later uploaded to the portal.

5. Steps 2–3 are repeated for unassigned students, treating the next school in their preference list that is not full as their "first preference", until either all students are assigned, all schools are filled, or there is no possible match.

The lottery-based allocation in Step 3 is central to our empirical strategy, which primarily compares lottery-winning students to lottery-losing students in oversubscribed schools. Causal interpretation of our treatment effect estimates relies solely on random allotment and not on the truthful revelation of preferences. This is important given that the Immediate Acceptance algorithm is not strategy proof, although we expect algorithm-related strategic considerations to be negligible in this setting.[17]

3 Effects on school choice and efficiency of public spending

3.1 Empirical strategy

We use the following specification to estimate the intent-to-treat (ITT) effect of being assigned a lottery seat:

$$Y_i = \alpha Z_i + \sum_x \gamma_x d_i(x) + v_i, \tag{1}$$

where Y_i indicates the outcome for child i and Z_i indicates winning the lottery for an RTE seat in a private school. This offer (Z_i) is randomly assigned conditional on applicants' ranking of schools but not unconditionally. Therefore, we condition on a vector of dummy variables $d_i(x)$ to account for the application choices of each student i ("randomization strata" or risk sets). Our coefficient of interest, α, is the ITT effect of being offered an RTE seat through the lottery.

Our preferred specifications adopt Abdulkadiroğlu et al. (2017)'s approach to controlling for applicant risk sets. We condition on a vector of narrow bins (of 0.001 probability each) of being assigned to a private school. We computed these probabilities by running 10,000 simulations of the assignment mechanism given the applicants' preferences. For each simulation, we recorded the school each student was assigned to. We then estimated, across all simulations, each child's probability of being assigned to a private school. The identifying assumption is that the offer of an RTE seat is conditionally exogenous after controlling for these narrow bins of the probability of

[17] The allocation rule was never mentioned in documents available to the public beyond stipulating that allocations would be lottery based. Further, as we discuss below, most students apply for only one private school. This makes sophisticated strategic considerations of how to order schools redundant. In 2020, partly reflecting feedback from our research team, the state moved to the (strategy proof) Random Serial Dictatorship mechanism to allocate seats.

an offer. For transparency and robustness, we also present estimates conditioning on the full vector of preference sets in Appendix E.[18]

Ex-post, some lottery losers may be assigned RTE seats in schools that still have space.[19] Since the policy variable is offering an RTE seat, we estimate, and focus on, the local average treatment effect (LATE) of being allocated an RTE seat.[20] We estimate the LATE by instrumenting an RTE seat assignment with winning the lottery. Specifically, we estimate the following equations via two-stage least-squares:

$$T_i = \beta Z_i + \sum_x \gamma_x d_i(x) + u_i, \quad (2)$$

$$Y_i = \delta \widehat{T}_i + \sum_x \gamma_x d_i(x) + \varepsilon_i, \quad (3)$$

where T_i indicates being assigned an RTE seat, and everything else is as in Equation 1. Here, δ is the effect of securing an RTE seat (through any means) on the outcome.

We compare the LATE estimate to the control mean for compliers (Imbens and Rubin, 1997). Our approach to estimation is based on Abdulkadiroğlu et al. (2018). See Appendix B for details and Abadie (2002, 2003) for relevant theoretical results.

3.2 Data

We use data from three sources: (i) application data provided by parents in 2019; (ii) two rounds of phone-based survey data collected by the research team to study enrollment; and (iii) administrative data on school characteristics. We describe each of these sources below.

3.2.1 Application data

We obtained data for all eligible applications submitted in 2019 through the online allocation system to implement the RTE in Chhattisgarh. The data has parents' rankings over schools, the assigned school (if any), and limited household characteristics, including their phone numbers. Parents applied to schools in March–April 2019 and were notified of the school assignment in May. Enrollment decisions were finalized by July 2019, well before the COVID-19 shock to education systems in March 2020.

In 2019, valid applications were received from 54,676 eligible students, 7,217 of whom were not allotted an RTE seat (see Table A.1; Panel A). Nearly half (48%) of the applicants were

[18] This latter strategy is inefficient, as limiting comparisons to exact matches discards much of the available variation. Our results are similar in magnitude and statistical significance across both procedures.

[19] This can be through subsequent lottery rounds, as well as through local education authorities.

[20] Nearly everyone ($\sim$ 95%) who is offered a seat, takes it. Thus, in practice, there is little difference between estimating the LATE of being offered a seat, and enrolling in an RTE seat.

female and 56% lived in a rural area. More than 50% of applicants have only one school on their preference list, and 92% have at most three preferences.

For ∼69% of applicants, the allocation system does not provide variation in whether they are assigned to a private school.[21] Our primary data collection focused on the remainder of the sample (N=16,703), for which we have some identifying variation on the extensive margin. One-third of these students were left unallotted (see Table A.1, Panel B). This subsample has a similar proportion of girls and number of schools applied for as the full sample that includes all applicants. However, the subsample is more urban (since urban areas are more likely to have oversubscribed schools) and, relatedly, has a lower proportion of Scheduled Castes and Scheduled Tribes. There are 5,863 schools in the lottery, each with roughly 10 seats available on average, but with 15 students applying for a seat.[22]

3.2.2 Primary data from phone surveys

We conducted two rounds of phone-based surveys to collect primary data on schooling choices, socioeconomic background and educational inputs from treated and untreated students. We randomized the order in which we called households in both survey rounds. All households were required to provide a phone number in order to apply; thus, all applicants are part of our sampling frame for phone-based surveys.[23]

First, between August and September of 2020, we attempted to call all individuals with an ex-ante probability of less than one of being allotted a private school quota seat (see Table A.1, Panel B) using the phone numbers provided by parents on their applications. We collected information about which school the applicant eventually enrolled in for the 2019–20 school year, along with basic school characteristics (e.g., medium of instruction and fee level) and household characteristics (parental education and occupation). We made up to five attempts to reach each household and completed interviews with about 45% of the targeted households.

Between November 2020 and January 2021, we attempted to recontact all households interviewed in the first phone survey and completed interviews with 59% of them. This second round collected detailed information on household assets which we use to compare the socioeconomic status of applicants to those of the eligible population. Since the original

[21] Given applicants' preference ordering, applications by other parents, and the number of seats available in each school/grade, these applicants are allocated to *some* private school with certainty (even if the private school they end up in is stochastic).

[22] Figure A.2 provides the full distribution of the number of applications schools receive. The average (median) school is ranked in 15 (10) applications. Schools are more likely to have seats available in Nursery than in Grade 1 (see Table A.1, Panel C). Table A.1 provides further details on the characteristics of applicants and schools, and Table A.2 explores how application behavior varies by household characteristics.

[23] According to the National Family and Health Survey of 2019–2021, almost all urban households (97%) and most rural households (92%) in Chhattisgarh have a mobile phone.

sample focused on applicants with an ex-ante probability of less than one of being allotted a private school quota seat, we also attempted to interview a random sample of 1,203 applicants who had a probability of one of being assigned to a private school. Of these 1,203 students, 462 answered our phone survey.

3.2.3 Administrative data on school characteristics

We use the U-DISE (Unified District Information System for Education) database, an annual census of all recognized (public and private) schools in the country, which contains information on school enrollment, infrastructure, and staffing.[24] We use data from the 2017–2018 school year, the most recent for which data were available at the time of writing. We merge the U-DISE data with a separate data set on school fees (for non-quota students) for recognized private schools.

3.3 Validity of the research design

3.3.1 Balance

We test for balance of observed characteristics in the applicant data and both phone surveys. Table 1 reports the results using our preferred specification, which conditions on bins of the probability of being offered a private school seat as in Abdulkadiroğlu et al. (2017), for all three samples. Table E.1 presents the results conditioning instead on the full vector of unique preference lists. Conditional on strata fixed effects, we do not reject the equality of mean characteristics across lottery winners and losers in any sample.

3.3.2 Attrition

Attrition is moderately unbalanced across lottery winners and losers: conditioning on the lotteries, we are slightly more likely — by 2.1 percentage points (over a base of 45%) in the first round and by 2.8 percentage points (over a base of 26%) in the second round — to reach students who were offered a seat than those who were not (see the last row in Table 1). Survey non-response is driven by being provided inaccurate phone numbers or failing to obtain a response even after five attempts. Attrition is higher for households in rural areas and those belonging to Scheduled Castes and Scheduled Tribes (see Table A.3). We investigate the sensitivity of our results to using low differential-attrition strata and Lee (2009) bounds. Our main findings are robust to these corrections.

3.3.3 Non-compliance / First stage

We verify that winning the first lottery corresponds to an offer of a free seat. Nearly all lottery winners reported having been allotted a seat (95%) in the first phone survey, but

[24] The U-DISE dataset does not include unrecognized private schools — schools that are operating without license or authorization from the government (Kingdon, 2020). This is not relevant in our setting since, by necessity, the policy only applies to recognized private schools.

so do about 22% of lottery losers (Table 1). As mentioned above, non-compliance among lottery losers (i.e., "always-takers") is expected, since education authorities attempt to fill vacant seats after the lottery-based allocation (the data we use) with unmatched parents. Compliance rates are similar in magnitude across grades (see Table A.4). We focus on the LATE of being allocated an RTE seat, using the outcome of the lottery as an instrument.

3.4 Effects of receiving a quota seat on enrollment decisions

Receiving a free seat may allow some quota-eligible students, who may not be able to secure admission or pay fees, to enroll in schools they could not attend otherwise. This potential shift in enrollment choices is the primary channel of (potential) impact for the RTE quota seats, and may operate on both the extensive margin, moving students into private schools (from no schooling or public schools), and the intensive margin, changing which private school they attend. We estimate policy-induced shifts on both margins.

3.4.1 Extensive margin of (private) school enrollment

The 4–6 age group, when students apply for RTE quotas, is a period of transitioning into primary schooling from either preschool or non-enrollment. Unlike primary schooling, which is mandatory from 6 years of age, preschool enrollment is neither universal nor compulsory. Guidelines for the enrollment age are often loosely applied. Therefore, children in this age group may be enrolled in a government childcare center or the pre-primary section of a private school, or enrolled in Grade 1 in either a government or private primary school, or not be enrolled in any preschool/school. Thus, a movement into the private sector can be induced on multiple margins. We collapse these possibilities into three states — (a) enrolled in a private preschool or school, (b) enrolled in a government school, and (c) not enrolled — and study the effects of being offered an RTE seat on each of these margins separately for the 2019–20 academic year.[25]

We note three results. First, nearly all applicants who were assigned an RTE seat were enrolled in private schools in 2019–20. However, this translates to around a 25 percentage-point (p-value < 0.001) increase in the probability of private school enrollment, as over three-quarters of compliers who did *not* receive an RTE place were also enrolled in private schools (Table 2, Columns 1–4).[26] Thus, the pool of applicants seems to disproportionately consist of students who would have attended private school anyway.

[25] We do not distinguish between non-enrolled and government daycare centers (called *anganwadis*), because the latter provide very little early childhood stimulation in practice (see e.g., Ganimian et al. (2021)). Nor do we distinguish between pre-primary and primary grades in private schools, since they exist in the same schools and kindergarten (preschool) classes serve as feeder grades into primary schooling (Singh, 2014).

[26] Throughout this section, and in what follows, we discuss LATE estimates as the principal parameters of interest. We present the ITT estimates only for transparency and do not emphasize them in the text.

Second, applicants assigned an RTE seat were 19 percentage points (p-value < 0.001) more likely to be enrolled in *any* school in 2019–20 from a base of 83% among the compliers (Table 2, Columns 1–4).

Third, the extensive margin effect is concentrated in the two preschool grades (Nursery and Kindergarten) that precede formal schooling, shifting some students from home care to private preschools. Applicants to Nursery who are assigned an RTE seat were 26 percentage points more likely to be enrolled in any school and 29 percentage points more likely to be enrolled in a private school in 2019-20; in Kindergarten, this declines to 16 and 24 percentage points, respectively; in Grade 1, this declines further to 3.6 and 13 percentage points. Thus, the steady-state effect of being allotted an RTE seat is likely to be around a 13 percentage-point increase in the probability of attending private school (the estimated effect in Grade 1 in 2019–20, when nearly all children were enrolled in school).

This extensive margin effect is small in relation to the implicit policy assumption that eligible students would not be able to access private schools in the absence of the quota. It is also modest compared to context-specific empirical benchmarks. Muralidharan and Sundararaman (2015) show that providing (untargeted) vouchers boosted private sector enrollment by ~ 50 percentage points in rural Andhra Pradesh; if financial constraints bind more strongly for poorer households, a reasonable prior would be to expect untargeted vouchers to provide lower-bound estimates for the effect on private school attendance for vouchers targeted at disadvantaged groups.[27] Our estimated effect is, instead, about half the size in the overall sample and about a quarter the size for applicants in Grade 1.

These results are robust to using only strata with no attrition, to focusing on strata with low differential attrition, and to Lee (2009) bounds correcting for differential attrition (see Table A.5). Since applicants who were allotted a seat were nearly universally enrolled in private schools, receiving an RTE offer through lottery eliminates baseline gaps in private school attendance by parental education, caste and gender (see Table A.7). The absolute treatment effect in Grade 1 (the steady-state) is still relatively modest in all subgroups.

3.4.2 Changes in the characteristics of the schools attended

Modest effects on the extensive margin may still be consistent with larger effects on the intensive margin since quota seats may change *which* private school a child enrolls in.

As a summary measure of quota-induced movement, we first examine whether quota students attend more expensive schooling options. In this context, school fees are the unsubsidized market price paid by non-quota students (taken from administrative data). The median private school in our sample charges INR 5,650 per year ($\sim$USD 75). The

[27] Andhra Pradesh is a richer state than Chhattisgarh. If financial constraints bind more in poorer contexts, we would expect treatment effects on private school enrollment to be larger in our setting.

distribution of annual private school fees varies from INR 2,100 ($\sim$USD 28) at the 5th percentile to INR 18,000 ($\sim$USD 240) at the 95th percentile. Public schooling and non-enrollment are free options (i.e., have a market price of zero).

The schooling choices of applicants allotted an RTE seat have a market price that is INR 4,630 (p-value $<$ 0.001) higher, on average, over a base of INR 5,292 (see Panel B - Table 3, Column 1). This treatment effect reflects both extensive margin shifts from zero-fee options (public schools and non-enrollment) to private schooling and movements within the private sector. The effect falls from Nursery to Kindergarten/Grade 1 as more applicants without an RTE seat move from non-enrollment to fee-charging private schools. Among applicants to Grade 1, when nearly all children are enrolled in schooling, the effect on market price is INR 2,881 (p-value $<$ 0.001), over a base of INR 5,946.

Next, we examine whether characteristics of the schools attended change in response to receiving a quota seat (see Table 4). We focus on applicants to Grade 1, nearly all of whom are enrolled in formal schooling, to avoid confounding effects on school characteristics with those on the extensive margin on school enrollment. In this sample, applicants assigned an RTE seat are 8.6 percentage points (p-value 0.05) more likely to attend English-medium schools (from a base of 50%). This increase is significant because English-medium instruction is perceived to have large labor market returns (Azam et al., 2013), and the average causal effect of attending private schools on student learning also appears to be greatest in English (Muralidharan and Sundararaman, 2015; Singh, 2015).

As a caste-based desegregation initiative, however, the quota seems ineffective: the average child allocated a seat is not exposed to a different socio-economic mix of peers (as measured by the proportion of students from Scheduled Castes and Tribes) than they would be without an RTE seat. This is also true if we explore heterogeneity by caste group. Scheduled Caste students allotted a seat do not attend schools with a different proportion of Scheduled Caste students. Likewise, Scheduled Tribe students allotted a seat do not attend schools with a different proportion of Scheduled Tribe students (see Table A.9).

Finally, there are no discernible differences in the schools attended by children who receive an RTE seat in terms of infrastructure, and only small differences in enrollment and pupil-teacher ratios.[28]

[28] These results are not mechanically driven by similarity between government and private schools. Schools in the two sectors differ substantially, even within narrowly defined communities, in the caste composition of their student body, as well as other observed characteristics such as medium of instruction, enrollment, number of teachers, and facilities. We discuss this further in Section 5.2.

3.5 Effectiveness of public spending: Inframarginality and Incidence

The policy aims to enable students from quota-eligible groups to enroll in schools they would otherwise be unable to attend. Treatment effects on student enrollment choices inform policy effectiveness in this domain, but a fuller assessment must consider at least two further questions. First, how effective is the public spending on this program in facilitating school choice? Second, and relatedly, to what extent does the policy succeed in targeting the households in greatest need of support?

3.5.1 Inframarginality

The effectiveness of program spending for school choice depends largely on the degree to which the implied transfer from quota seats is inframarginal to household choices.

As a first step towards examining inframarginality, we ask what proportion of students would have attended their top-choice school even without getting an RTE seat *at that school*. Our specification is analogous to that used to estimate the intent-to-treat in Table 2, except that the treatment (lottery-based RTE offer) is specific to the top-choice school rather than any school.[29] We find that 39% of students who did not receive a lottery-based offer of a seat at their top-choice school are nonetheless enrolled in their top choice; this figure rises by 57 percentage points for students who were offered an RTE seat (Table 5, Column 1). In Grade 1, our steady-state sample, 50% of students who lost the lottery for their top-choice school, enroll in it anyway; winning the lottery raises this figure by 43 percentage points. Overall, quota seats are completely inframarginal to the choices of a substantial share of recipients (although many students also use it to upgrade to a more-preferred school).[30]

To consider the effectiveness of program spending, we ask what proportion of *expenditure* on quota seats is inframarginal. Our benchmark here is the causal effect of receiving a quota seat on the market price of the schooling option (i.e., the average economic value of the improvement in educational options received by beneficiaries). This sum, which is INR 2,881 in Grade 1, represents the lowest mean value of a top-up voucher required for parents to choose the same options as they avail in the quota regime. This thought experiment, which is infeasible because parents' willingness-to-pay for schooling options is not observed, takes the pool of applicants, their preferences, and the availability of seats as given. This estimate, first reported in Table 3, is repeated in Table 6 for convenience.

[29] We do this to avoid violations of monotonicity. Although the offer of a free quota seat in their top-choice school makes a student more likely to attend that school, an offer for their second-choice school may make her *less* likely to enroll in the top-choice school as a fee-paying student. We report evidence of such cross-partial effects when regressing enrollment in the top-choice school on a vector of offers at top/second/third schools (Table A.8).

[30] As mentioned in Section 2.3, we expect parents' first-choice school to reasonably reflect their true preferences in this setting, even though the Immediate Acceptance mechanism is *not* strategy-proof.

We compare the average cost of a quota seat for the government to this benchmark.[31] This sum is given by the fees charged by the allotted school up to a maximum of INR 7,000 (Table 6, Panel B). In Grade 1, this sum averages INR 5,795. Thus, in Grade 1, approximately half of the reimbursed amount is inframarginal to school choice.[32]

However, the total cost of the program must also take into account private schools' contributions — 30% of schools charge a higher fee than the reimbursement cap of INR 7,000.[33] The total cost is identical to estimating reimbursements without the capped limit of INR 7,000. This sum is INR 8,826 on average, which is $\sim$3 times the incremental educational expenditure on fees received by the beneficiaries. The difference between the "full cost" and the reimbursed value of the RTE seat effectively represents a tax on high-fee private schools (imposed by the cap). This effective tax may partly explain the strong opposition to this policy by elite private schools in many states across the country. The value of the tax is similar to the net incremental value in school fees that students gain. In summary, at least for applicants who apply to oversubscribed schools, a substantial portion of the average cost of the quota — about 50% of the reimbursed amount and 67% of the total cost — is inframarginal to school choices.

3.5.2 Incidence of policy benefits

The modest extensive margin effects and the high degree of inframarginality suggest substantial selection in the pool of applicants (and, thereby, in the pool of quota recipients). For example, our estimated control complier mean of 78% is substantially higher than the statewide average of 27% of Scheduled Caste students in Grades 1-3 who attend a private school (computed based on U-DISE data for 2017-18).

A potential concern for this interpretation is that our estimates of counterfactual enrollment and inframarginality are based on applicants subject to lottery-based allocation. This sample may not be representative of the overall pool of applicants. We explore this by comparing the observable characteristics of *all* quota applicants, as measured in our second phone survey, which also included a random sample of applicants to undersubscribed

[31] This exercise ignores income effects from the transfer (treating them as small in relation to annual household budgets). We also disregard the welfare effects of the inframarginal portion of the expenditure, which is effectively a cash transfer, as these are outside the policy objectives. In this, we follow the long literature on the impact of educational vouchers and other inputs in multiple settings (Epple et al., 2017).

[32] For students who would have attended the *same* school in the absence of quota, the entire expenditure is inframarginal. For many other students, some of the expenditure is still inframarginal — the difference between the fees reimbursed by the state and what they would have paid in school fees (potentially in a different school) in the absence of the policy.

[33] We disregard any savings that the government might make from students shifting out from the government sector since, in the short-to-medium term, the overall costs are fixed. In particular, the government cannot adjust the wage bill for permanent teachers (who account for the bulk of recurring costs).

schools, to all households in Chhattisgarh using representative data from the National Family Health Survey (NFHS) from 2019–2021.[34]

We restrict the NFHS sample to households with children aged 4–7 and present estimates for the overall population and individual caste groups. We separately present the observed characteristics of (i) lottery-based applicants ("stochastic allocation"), (ii) other applicants ("deterministic allocation"), (iii) all applicants, and (iv) the NFHS representative data. We focus on two markers of socioeconomic advantage. The first is asset ownership, which we summarize with an index based on a Principal Component Analysis (Filmer and Pritchett, 2001) (see Table A.11 for detailed asset information). The second margin is maternal and paternal education, which we summarize as whether the parents have above primary education or not (see Tables A.12 and A.13 for detailed parental education information). Both asset ownership and parental education were elicited in our survey using the same questions as the NFHS. We generated the asset index comparably across datasets by running the Principal Components Analysis on an appended dataset that included all observations across the two sources (ensuring that the same factor loadings were applied).

Overall, the average applicant is more likely to have parents who own their own house, and have more assets and education than the average child in the state, even without conditioning on eligibility for an RTE seat (Panel A, Table 7). These differences are particularly stark in the case of parental education. Applicants are also better off within each caste group (Panels B-D, Table 7). Thus, the policy benefits seem to accrue largely to socioeconomically-advantaged members of quota-eligible groups. Although significant, differences between the lottery-based and deterministic applicants (Column 5, Table 7) are typically much smaller than the differences between all applicants and the statewide data (Column 6, Table 7).

Finally, we map quota applicants to representative data to estimate the proportion of applicants and quota-recipients in each quintile of the state-wide socioeconomic distribution of households with children aged 4–7 years. Application rates — and therefore the allocation of seats and public expenditure — are regressively distributed. 25.2% of applicants (and RTE seat holders) come from the top quintile of the economic distribution, compared to 8.0% from the bottom quintile (see Figure 1a). These figures do not condition on caste or socioeconomic status: thus, quintiles of the socioeconomic distribution refer to the *full* statewide distribution for households with eligible aged children. As a consequence of selection into applications, 24.3% of the monetary benefits of this policy (i.e., the fee of the schools for those admitted to under the quota) accrue to

[34]We use inverse probability weights to account for differential probability of being surveyed and differential non-response to the phone survey (using household characteristics to predict the likelihood of having completed a survey).

children in the top quintile of the SES distribution, compared 7.4% to children in the bottom quintile. Students in the top quintiles are also more likely to apply to schools with fees above the reimbursement cap — thus, receiving larger subsidies from schools (see Figure 1).

4 Why are application rates low for poorer households?

Receiving quota seats has limited effects on improving access to better private schools primarily because poorer quota-eligible students do not apply. Understanding the reasons for this regressive selection into quota applications is, thus, central for identifying ways to improve policy effectiveness. In this section, we focus on household preferences, information constraints and application frictions as potential explanations. We explore spatial supply-side constraints (i.e., the availability of nearby private schools) in Section 5.2.

4.1 Data

Investigating parental preferences and information sets requires representative data on quota-eligible students, not just those who applied. We collected this in urban and rural areas of Raipur, which is the most populous district of the state. This district accounts for ∼8.4% of the state's population, 15.9% of the population of private school students and 12% of the total eligible applicants for RTE seats in 2019. It also includes the state capital. An estimated 43–47% of the population of the district was below the official poverty line (World Bank, 2016), 16.6% belong to Scheduled Castes and 4.3% belong to Scheduled Tribes (Government of India, 2011).

To draw our sample, we first selected a random set of 20 locations each in urban and rural areas of the district.[35] In these locations, we interviewed all households that had a child aged between 3–5 years of age (N=1,059, identified based on a door-to-door listing of 12,225 households).[36] These households were administered an extensive questionnaire that we designed to directly study how parental preferences and application constraints vary across the SES distribution, as described below. We measure socioeconomic status using an index created from household ownership of assets, consumer durables, and housing quality using Principal Components Analysis (Filmer and Pritchett, 2001).

[35] We drew a sample of squares on Google Maps of 1km x 1km in rural areas and 300m x 300m in urban areas. After verifying that these included habitations (e.g., excluding exclusively agricultural land), we identified a government early childhood care center (anganwadi) towards the center of the square, and re-centered the square around it.

[36] Our sample is similar in terms of socioeconomic markers to the population of Raipur (see Table A.15).

4.2 Demand for private schools from low-SES households

Regressive selection is inevitable if too few poor households (compared to the number of seats available) wish to send their children to private schools even for free.[37] Similarly, there is limited scope for reducing inframarginality if only a few households want to attend private schools but are financially constrained from doing so.

To understand preferences over different schools, we adapt the methodology of Delavande and Zafar (2019). Specifically, we provide parents with five (fictitious) schools, representing the range of choices available in similar markets, from which they must choose where to enroll their child. Each school is characterized by a vector of characteristics, including the fees they charge. We ask parents to rank the schools from 1 (most preferred) to 5 (least preferred) in three scenarios: (a) when only public schools are free, while private schools charge posted tuition fees ("status quo"), (b) when one private school, randomly-chosen, is made free through a school-specific tuition waiver ("RTE quota") and (c) when all schools are made free through an unconditional scholarship to the student ("voucher"). We randomize the sequence in which the households answer the scenarios — {status quo, RTE quota, voucher} and {voucher, RTE quota, status quo} — to avoid framing effects.

We take the bundle of characteristics of each school from the distribution of schools in Raipur district in official data. Specifically, we provide information on the number of classrooms, total enrollment, the number of teachers, location (distance to the household), the highest class offered, and the unsubsidized fee levels. Two schools are public: the first is "nearby" and "small" (distance is randomly assigned between 250 to 750 meters, and other characteristics at the 25th percentile of public schools by enrollment); the other public school is "big" (benchmarked to characteristics at the 75th percentile of public schools) and typically "further away".[38] The other three schools are private, with characteristics corresponding to the median school in each tercile of the fee distribution, and the fees are randomly assigned from the approximate range in each tercile. Distance to each private school is randomly assigned from 0.5 to 2 km. Importantly, reflecting the local setting, we do not provide any information on student achievement in schools: in this context, there are no standardized exams in primary schools and parents do not have access to any ranking of primary schools.[39] Thus, the characteristics on which we provide information are the relevant characteristics that can be measured for the entire market.

[37] Low demand may result if, for example, low-SES parents expect that private schools will not adequately cater to their children's learning needs or that they will face discrimination from other students or teachers. Bau (2022) documents that private schools in Pakistan have a higher value-added for richer students.

[38] Benchmarks are generated separately for urban and rural locations. Please see Appendix C for details on design, instruments and data validation.

[39] No reliable public data source measures student achievement in all private and government schools (or in representative samples) in India overall or this state. The most widely-used nationwide independent source, Pratham's ASER surveys, only covers rural areas and does not provide any school rankings.

Our principal goal is to study the overall demand for private schools by socioeconomic status and the extent of financial constraints. We are mainly interested in two quantities and how they vary across the SES distribution. First, the proportion of households who choose a private school as their top choice in the "voucher" scenario. This provides a direct measure of whether too few poor parents demand private schools, even when offered for free without selective admissions. Second, we want to compare this quantity to responses by the same parents when they have to pay the market price for private schools. This provides an estimate of the proportion of parents across the SES distribution who are financially constrained in being able to send their children to private schools, i.e., the group for whom vouchers may induce extensive margin shifts to private schooling (with zero inframarginality of spending).[40]

We plot these quantities non-parametrically in Figure 2, which shows three important results. First, low parental demand for private schooling is not a binding constraint to applying for quota seats: when offered for free, the majority of parents across the SES distribution prefer a private school as their top choice and, in the bottom quintile of the SES distribution, this figure is 53%. Second, the policy is not misguided in perceiving financial constraints as a major barrier to private school access for disadvantaged households: demand for private schools is much lower in the "status quo" scenario than the "voucher" scenario, especially for poorer households. Approximately 40% of all below-median SES households appear financially constrained on this extensive margin. Third, there *is* an SES gradient in the demand for private schools, which is steeper for below-median households. However, it cannot explain undersubscription by itself because the proportion of poor households is large and base demand is high.

We further elicited what parents expect their child's experience to be in different schools. Specifically, we asked parents on a 5-point scale to report how likely it was that their child will make friends, that the teacher will pay attention, that the child will get a good job at age 30, and that the child will be happy in the school.[41] In all dimensions, parents expect their children's experience in private schools to be better than in the nearby government school (Figure 3). Overall, low demand among poorer households appears unlikely to fully account for the low application rates from these households.

[40] Although we focus on only these two quantities in the main paper, our hypothetical choice data can provide substantial further insight into take-up and the demand for specific private schools. In Appendix C, we use parental choices in the "RTE quota" scenario to further characterize the proportion of inframarginal take-up across the socioeconomic distribution. The principal insight from this exercise is to directly link inframarginality of program spending to program take-up across the SES distribution. The data can further be used to estimate random utility models to study parental demand for specific school attributes (like Delavande and Zafar (2019) and Wiswall and Zafar (2018)), an exercise we intend to pursue in further work. We forego the estimation of the full model here given our principal interest in studying the overall demand for private schools and its variation over the socioeconomic distribution.

[41] To reduce survey fatigue, we only asked this for one government and one private school.

4.3 Survey evidence on information constraints and ability to apply

Financially-constrained eligible households may not avail RTE quotas, despite valuing private schooling, because they did not know about the policy or how to apply. They could also face additional frictions in the application process.[42]

Our household survey data suggests such constraints are important in this setting (Figure 4). First, information constraints seem to be both large and regressive. There is a stark SES gradient in whether households, with children between 3–7 years, have even heard of the existence of the policy: 65% of parents in the top decile have heard of the policy, while only 20% of parents in the bottom decile report the same.[43]

Second, potential frictions to applying also seem to differentially burden poorer households. While nearly all (98%) high socioeconomic status households have access to internet, only 9.1% of households in the bottom decile have internet connection (see Figure 4). This differentially burdens poorer households from accessing the centralized application portal: although households can apply through internet cafes or government service centers if they do not have internet access at home, these alternatives are more burdensome than applying at home. Fewer low-SES households report having ever applied for a government benefit online and fewer of them can access support from others for doing so.

4.4 Experimental evidence on addressing application frictions

Our survey evidence suggests that information and application frictions constrain potential applicants. Yet, if further constraints bind, relaxing these specific constraints alone may be inadequate to boost application rates or reduce regressivity. We investigate this through a randomized field experiment — in the same sample as our household survey — to evaluate an intervention to relax information frictions and application complexity. Specifically, by actively assisting households in applying, we aim to understand (i) constraints that may remain even after dealing with (apparent) first-order frictions and (ii) the extent to which this intervention might help address issues of undersubscription and regressive selection.

4.4.1 Intervention and Experiment Design

Our intervention aimed to deal with the twin constraints facing households of insufficient information about the policy and the lack of means to apply using online portals. We

[42] Such administrative burdens are consistent with evidence on the positive effect of information campaigns on the take-up of welfare programs and student aid. See, e.g., Bettinger et al. (2012); Hoxby and Turner (2013); Bhargava and Manoli (2015); Neilson et al. (2019). See, also, the theoretical framework in Finkelstein and Notowidigdo (2019) on how such differential burdens may explain take up for a range of public policies.

[43] This result is robust to focusing on heterogeneity within rural/urban areas, and focusing on heterogeneity within each sampling unit (i.e., small geographical areas of 9 hectares in urban areas and 100 hectares in rural areas) — see Table A.14 and Figure A.3.

focus on both constraints given evidence from other sectors and countries that information provision alone may not suffice to raise take-up when households also face other barriers (Banerjee et al., 2021; Finkelstein and Notowidigdo, 2019).

We used our in-person household survey from February 2022 as a baseline and restrict our attention to households that reported having a mobile phone (N=914, out of 1,059 households in total). This sample was not selected based on quota eligibility (or predictors thereof). Thus, our estimates are informative about the effectiveness of untargeted interventions in this setting (which are common and reflect the difficulty for external organizations, like NGOs, to verify eligibility ex ante). We randomized ∼50% of these households in each locality (N=459) to be offered the intervention. Observed characteristics (including previous knowledge of the RTE policy) are balanced between the treatment and the control group (see Table 8).

The intervention consisted of providing detailed information about the policy and, if the household was willing to apply and eligible, to assist them in applying. The intervention was implemented between April 15 and May 2, 2022, when quota applications were open, and consisted of the following, sequential, steps:

1. Call all treatment households to offer information and potential support to apply for a quota seat (if eligible). If households could not be reached on the phone number they provided, we revisited the household to elicit this information.
2. Households were asked if they had heard about the RTE quota, whether they knew that the applications were currently open, whether they had thought already of applying to the program, and if they were interested in receiving more information.
3. Interested households were provided detailed information on the eligibility criteria for the policy and the documents required for demonstrating eligibility. Surveyors collected information on each of the documents that the household reported having.
4. If a household reported having all documents, we offered application support and made an appointment to visit the household. If a household did not have all the documents, we provided information on where they could obtain the necessary documentation. Applicants were provided a number that they could call if they succeeded in getting these documents to receive further support.
5. For interested households (which reported having all documents), an interviewer visited them at home to help them fill out the application online.[44] They filled out identifying details, uploaded the documents, showed the households the list of schools available in the portal for their area, and ranked the schools as directed by the

[44]If the household did not have required documents during the home visit, surveyors revisited them after a week. If the household didn't have the necessary documents even on the second visit, we provided them with a number that they could call if they succeeded in getting these documents to receive further support.

parents. If parents wanted, they could submit the forms immediately. If they wanted more time to think about it, the surveyor saved the form and provided the household with the relevant login details so they could submit applications themselves later.

This intervention design resembles common interventions in this setting and is thus a natural benchmark for future interventions. Our focus, as with prior experiments evaluating take-up, is on the rate of successful applications.

Given stark SES gradients in information and internet availability in the survey, we expected that treatment effects would be strongly progressive (i.e., that the intervention would raise application rates more for poorer households). The intervention itself was not, however, targeted by socioeconomic status or caste. This design choice reflects two main considerations. First, SES information at the household level is not typically available to civil society organizations (or even government bodies) in this context; thus, targeting on individual characteristics is not typically a feasible policy choice. Second, for our primary aim to isolate barriers to application and how they vary across the socioeconomic distribution, it is useful to have uniform coverage across SES. In particular, uniform coverage over the full distribution helps diagnose the extent to which targeting (feasible only for limited characteristics and costly) may be needed at all.[45]

4.4.2 Implementation and take-up of application support

Our first set of insights comes from the *process* of delivering the intervention.

Of 459 households randomized into treatment, we were unable to reach 21% (N=95) in the short application/treatment window despite multiple contact attempts by phone and home visits. Confirming high demand for private schooling, very few of the remaining households (N=36) declined the offer of further information and support. A further 17% of households (N=80) reported having already filled out the form, including in previous years.

The main constraint in submitting applications was the availability of documents to certify eligibility. Of the remaining 248 households, 186 did not have the full set of documents (mostly proof of being income-poor or from a disadvantaged caste). This group conflates both those who are genuinely ineligible for the quota and those who are eligible but cannot document it. However, the correlation between having had documents or having already applied and SES is instructive: richer households are much *more* likely to have documents certifying eligibility than poorer households (see Figure 5). In the bottom decile of the SES distribution in our sample, about 24% of households can document eligibility; this figure rises to about 56% for the top decile. Finally, 19 households did not have a private school listed in their neighborhood.

[45]We discuss the potential for targeting to improve interventions further in Section 5.

We eventually helped 9.4% of households that were randomly assigned to be treated submit a form (43 out of 459 households randomized into treatment).

4.4.3 Treatment effect on application rates

We attempted to re-interview all households in our experimental sample after the application window closed to elicit whether they applied for an RTE seat (our primary outcome), whether they secured a seat, and whether they used an RTE seat to enroll their children in school. We were able to reach ∼80% of them, without any differential attrition between the treatment and the control group (see Table 8).

The intervention boosted application rates by 9.5 percentage points (p-value .0037), which is a 43% increase over a control mean of 22% (see Table 9), a large effect in relative terms.[46] The increase in application rates is nearly identical to the proportion of randomly-assigned households whom we were able to help submit applications, indicating that our intervention did not crowd out alternative sources of application help. Knowledge of the existence of the policy is very high in the control group since we had asked about the policy to all households in the sample in the previous month (and hence they had heard about it).

This increase in application rates also translates to 3.3 percentage point increase in the probability of being allocated an RTE seat (an increase of 40% over a control probability of 8.2%), although we lack sufficient statistical power to detect this effect (p-value .12).[47] Despite these positive results, however, our experiment also indicate the limits of such an intervention. Supporting households at the time of application, as is common for many interventions, does not allow them enough time to obtain documents certifying eligibility — this limits the prospects of increasing take-up among eligible households.

Constraints posed by the availability of documents — which our intervention did not address — also affects the distribution of treatment effects. Contrary to our priors, the intervention did not differentially benefit poorer households (see Figure 6): we are unable to reject equality of treatment effects across terciles of socioeconomic status and the point estimates also appear very similar.[48] Overall application rates in our treatment group by quintiles of SES closely mirror the gradient in the availability of documents in Figure 5. This pattern is intuitive, given that our application support did not crowd out other ways of applying, and suggests that the lack of documentation was a binding constraint.

[46] All treatment effect estimates are intention-to-treat effects since not all households that were randomized into treatment could be contacted during the application window. Results are robust to controlling for household socio-demographic characteristics, as well as preferences for private schools.

[47] The effect on the probability of being allocated an RTE seat is aligned with the effect on the application rate, once we take into account that an application in the control group has a ∼37% probability of translating into an offer for an RTE seat.

[48] We confirm the equality of treatment effects across SES at other quantiles of socioeconomic status in Table A.16.

5 Discussion

5.1 Multiple constraints and implications for policy design

Our survey evidence and the effects of our randomized intervention both highlight the importance of multiple constraints. All constraints must be satisfied to apply for quota seats, as in an "O-ring" process (Kremer, 1993).[49] Since all constraints we examined affect poor households more, this has important implications for policy design.

That multiple constraints, and especially those resulting from administrative burdens, may limit the overall effects of application assistance is suggested by, e.g., Banerjee et al. (2021). However, we show that this not only affects average treatment effects, but also the *composition* of who is affected. Specifically, in a setting with multiple constraints, who is moved by the intervention into applying successfully is determined by the incidence of constraints left unresolved by the intervention, not just the constraints that it successfully resolved. Treatment effects from our intervention are not progressive, despite both information and application complexity disproportionately affecting the poor, because the unresolved constraint (documentation) *also* disproportionately constrains poorer households.[50]

This broad insight also translates into practical recommendations for future interventions in this setting. First, interventions would ideally provide information months before application deadlines to allow households to procure documents. Further, this would need to be complemented by assisting households to apply in later stages.

Second, partial interventions that only relax *some* constraints may worsen regressive selection. Since common interventions — such as text messages or the introduction of helplines and chatbots for application assistance — are often easier for less-poor households to access and respond to, regressive selection is hard to remedy using current common low-touch interventions. This is particularly relevant for policies where the good being provided is valued as much or more by the less-poor (like private school quota seats).

Third, targeting these supplementary interventions may have very high returns for improving the composition of beneficiaries, even where overall policy eligibility rules are hard to change. Such targeting could be geographic — e.g., targeting areas with high

[49]The existence of an "O-ring" process can be inferred by the strong dependence of treatment effects on having requisite documentation: without documentation, no willing households could complete applications; with documentation, nearly all of them could (except for a subset further constrained by school availability in their neighborhood).

[50]Since multiple constraints may apply, all must be satisfied for payoffs to be realized. Consequently, the effects of individual constraints are *not* additive and we cannot decompose the non-applying eligible population across multiple possible constraints such as information, the ability to access online portals, or the ability to provide documentation.

poverty rates where quota seats exist and are undersubscribed. It could also be based on within-area proxies that are easily observed — e.g., targeting the parents of children who currently attend government preschools may substantially reduce inframarginality (since these children are more likely to attend government primary schools). Optimal methods for targeting interventions and information are an active area of research in both development economics and public economics (see, e.g., Alatas et al. (2012, 2016); Banerjee et al. (2018, 2019)). These are likely to be particularly relevant also for redistributive policies such as the private school quotas. Improving the incidence of quota benefits is desirable both on equity grounds and also for reducing inframarginality in program spending and, therefore, policy effectiveness.

Although default interventions are less promising in our setting, better interventions only require combining elements that are already common. In India (and elsewhere), multiple interventions exist to (separately) provide information about schemes, help households get documentation, and help households apply; combining these is logistically feasible and desirable. The development of such multi-component programs follows directly from evidence on the complementarity of different program components in policy design (Mbiti et al., 2019) and resembles bundled interventions in several sectors (see, e.g., Banerjee et al. (2015); Bandiera et al. (2020); Muralidharan et al. (2019)).

5.2 Spatial segregation and the limits to quota-induced reallocation

Our analysis mainly focused on understanding why households do not apply for quota seats even if private schools are available. This focus was appropriate because information constraints and application hurdles are relevant margins for policy action to improve take-up. For the policy overall, however, a further constraint to *overall* effectiveness for school integration is whether private schools are located in communities with high proportions of disadvantaged households. If most quota-eligible households have few private schools available in their communities, the policy would be limited in the ability to reduce stratification (see, e.g., Monarrez (2022); Campos and Kearns (2022)).

In our final analyses, we study the extent to which such spatial constraints might matter. To this end, we assembled geolocated data on population and poverty rates for individual villages and towns (Asher et al., 2021), with GPS locations of recognized schools and administrative data on school characteristics.[51] Private schools are common in rural and urban areas, but not universal. They are less likely to be available in communities with a higher proportion of quota-eligible disadvantaged groups (Figure

[51]We retrieved school GPS locations from the official website https://schoolgis.nic.in/ in October of 2021, which we matched to school management and enrollment data from U-DISE. We matched 99.8% of students in government schools and 96.3% of students in private schools to their locations.

7). In the state overall, 43.4% of Grade 1 enrollment is in communities with only public schools; this figure is 53.8% for SC/ST students and 55.7% for income-poor students.[52] Thus, the spatial distribution of schools and disadvantaged groups *is* relevant for explaining state-wide differences in the probability of private school enrollment for disadvantaged and non-disadvantaged households.

However, communities that have both private and public schools still show considerable sorting by caste across schools. The proportion of students from Scheduled Castes and Tribes is 28.0 percentage points higher in public schools state-wide — 45.7% of this difference (12.8 percentage points) is within communities. Private and government schools also vary substantially in size, infrastructure, staffing, and medium of instruction within the same communities (see Table 10). Thus, although spatial constraints limit the equalization of caste composition in even the best-case scenario, it appears that substantial reductions in segregation remain possible. We confirm this intuition in Appendix D by simulating potential reallocation of RTE quota seats within postcode or a proxy for "neighborhood".[53] Reallocation of quota seats can reduce the public-private difference in the proportion of SC/ST students by half (to $\sim$17 percentage points).

6 Conclusions

Quotas in private schools are the main policy vehicle used to address educational segregation in Indian primary schools, of which private schools form a substantial share. We evaluated whether the policy, as implemented in Chhattisgarh, delivers on that promise.

Our results paint a complex picture. Conditioning on the set of applicants, the policy delivers large monetary gains to applicants who receive a free place. Obtaining a seat induces some quota-eligible students to attend preschool and others to attend schools they would not have been able to afford. Yet, this success is qualified. The quota is used primarily by households that would send their children to private schools anyway and much of the expenditure is inframarginal to school choice.[54] The policy, thus, largely acts as a transfer for beneficiaries without achieving the goal of changing the composition of classrooms by admitting financially constrained disadvantaged students.

[52]There is no direct poverty measure in the U-DISE data. We assume the proportion of enrolled children who are income-poor in a given neighborhood matches the proportion of income-poor households. This is a conservative estimate since fertility is typically larger in poorer households (IIPS, 2017).

[53]We approximate neighborhood using the SHRUG-ID created by Asher et al. (2021).

[54]A useful comparison for our results is the evaluation of the PACES voucher scheme by Angrist et al. (2002) in Colombia. Like us, they find modest treatment effects on the extensive margin of private school enrollment ($\sim$15%). The PACES program, however, required applicants to have sought and secured admission to a private school before applying. While the RTE quota did not feature this requirement, selection of a similar magnitude seems to have occurred *de facto*, subverting the explicit policy goal of expanding access to private schools for disadvantaged groups.

We show that these policy effects on school choice mainly reflect selection into applications: poorer students are less likely to apply, even when eligible. This does not primarily reflect low demand but rather a succession of constraints that differentially impact poorer households more. These constraints — information, application complexity, and documentation — are addressable by policy interventions. Given the substantial proportion of poorer households that would like to enroll children in private schools but cannot afford to (Figure 2), our results suggest substantial scope for improving policy effectiveness even within current policy rules. To address the regressive selection, however, potential interventions will need to consider the joint incidence of multiple constraints; optimal interventions to address regressive selection are likely to differ from those focused only on improving application rates, both in design and in targeting. These insights may be relevant for the take-up of redistributive programs more generally beyond this policy.

Our results from Chhattisgarh are likely informative about the RTE policy across India. RTE quota seats are undersubscribed in all states in India.[55]. It is unlikely that this can be fully rationalized by low demand from eligible parents for private schooling. Instead, the information and application constraints we document are likely to be important across settings. The only other evidence on extensive margin effects documents similarly small effects of receiving RTE seats on private school attendance in Grade 1 in a much richer state (Damera (2017) in Karnataka) indicating that concerns of inframarginality are also likely to be much broader than Chhattisgarh alone.[56]

Finally, we have focused solely on policy effects on enrollment and their incidence. Understanding downstream policy effects on, for example, the social integration of quota-admitted students in classrooms, learning outcomes, non-cognitive skills, and, eventually, effects in adulthood should be an area of priority for further research.

References

ABADIE, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284–292.

——— (2003): "Semiparametric instrumental variable estimation of treatment response models," *Journal of Econometrics*, 113, 231–263.

[55]Fill rates for RTE seats in 2018-19 ranged from 7% in Uttar Pradesh to 70% in Rajasthan, with other major states like Madhya Pradesh (43%), Delhi (56%) and Tamil Nadu (60%) in between (Indus Action, 2021).

[56]That being said, there is substantial variation across states in aspects of policy implementation. This includes when states implemented the reform, rules for eligibility, the definition of "neighborhoods", and application and admissions processes. States differ also in the baseline spread of private schooling, both spatially and as a share of enrollment. Building an evidence base spanning multiple heterogeneous states could inform better policy design as in, for example, the substantial literature on school choice mechanisms (Pathak, 2017), voucher systems (Urquiola, 2016) and Charter Schools (Cohodes and Parham, 2021).

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017): "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation," *Econometrica*, 85, 1373–1432.

Abdulkadiroğlu, A., P. A. Pathak, and C. R. Walters (2018): "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economics*, 10, 175–206.

Ahern, K. R. and M. Giacoletti (2022): "Robbing Peter to pay Paul? The redistribution of wealth caused by rent control," Tech. rep., National Bureau of Economic Research.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, and J. Tobias (2012): "Targeting the poor: evidence from a field experiment in Indonesia," *American Economic Review*, 102, 1206–40.

Alatas, V., R. Purnamasari, M. Wai-Poi, A. Banerjee, B. A. Olken, and R. Hanna (2016): "Self-targeting: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 124, 371–427.

Andrabi, T., N. Bau, J. Das, and A. I. Khwaja (2022): "Heterogeneity in School Value-Added and the Private Premium," Working Paper 30627, National Bureau of Economic Research.

Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer (2002): "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment," *American Economic Review*, 92, 1535–1558.

Arcidiacono, P., V. J. Hotz, and S. Kang (2012): "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 166, 3–16.

Arcidiacono, P. and M. Lovenheim (2016): "Affirmative action and the quality-fit trade-off," *Journal of Economic Literature*, 54, 3–51.

Asher, S., T. Lunt, R. Matsuura, and P. Novosad (2021): "Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in India using the shrug open data platform," *The World Bank Economic Review*, 35, 845–871.

Attanasio, O. P. and K. M. Kaufmann (2014): "Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender," *Journal of Development Economics*, 109, 203–216.

Azam, M., A. Chin, and N. Prakash (2013): "The returns to English-language skills in India," *Economic Development and Cultural Change*, 61, 335–367.

Bagde, S., D. Epple, and L. Taylor (2016): "Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India," *American Economic Review*, 106, 1495–1521.

——— (2022): "The emergence of private high schools in India: The impact of public-private competition on public school students," *Journal of Public Economics*, 215, 104749.

Bandiera, O., N. Buehren, R. Burgess, M. Goldstein, S. Gulesci, I. Rasul, and M. Sulaiman (2020): "Women's empowerment in action: evidence from a randomized control trial in Africa," *American Economic Journal: Applied Economics*, 12, 210–59.

Banerjee, A., E. Breza, A. G. Chandrasekhar, and B. Golub (2018): "When less is more: Experimental evidence on information delivery during India's demonetization," Tech. rep., National Bureau of Economic Research.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019): "Using gossips to spread information: Theory and evidence from two randomized controlled trials," *The Review of Economic Studies*, 86, 2453–2490.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry (2015): "A multifaceted program causes lasting progress for the very poor: Evidence from six countries," *Science*, 348, 1260799.

Banerjee, A., A. Finkelstein, R. Hanna, B. A. Olken, A. Ornaghi, and S. Sumarto (2021): "The challenges of universal health insurance in developing countries: experimental evidence from Indonesia's National Health Insurance," *American Economic Review*, 111, 3035–3063.

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

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012): "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India," *science*, 335, 582–586.

Bergman, P. (2018): "The Risks and Benefits of School Integration for Participating Students: Evidence from a Randomized Desegregation Program," Tech. rep., Institute of Labor Economics (IZA).

Bertrand, M., R. Hanna, and S. Mullainathan (2010): "Affirmative action in education: Evidence from engineering college admissions in India," *Journal of Public Economics*, 94, 16–29.

Besley, T. and S. Coate (1992): "Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs," *The American Economic Review*, 82, 249–261.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012): "The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment," *The Quarterly Journal of Economics*, 127, 1205–1242.

Bhargava, S. and D. Manoli (2015): "Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment," *American Economic Review*, 105, 3489–3529.

Bhavnani, R. R. (2017): "Do the effects of temporary ethnic group quotas persist? Evidence from India," *American Economic Journal: Applied Economics*, 9, 105–123.

Bleemer, Z. (2022): "Affirmative action, mismatch, and economic mobility after California's Proposition 209," *The Quarterly Journal of Economics*, 137, 115–160.

Boneva, T. and C. Rauh (2017): "Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns," CESifo Working Paper Series 6756, CESifo.

Campos, C. and C. Kearns (2022): "The Impact of Neighborhood School Choice: Evidence from Los Angeles' Zones of Choice," *Available at SSRN 3830628*.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng (2019): "On binscatter," *arXiv preprint arXiv:1902.09608*.

Chattopadhyay, R. and E. Duflo (2004): "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India," *Econometrica*, 72, 1409–1443.

Clotfelter, C. T. (2011): *After Brown: The rise and retreat of school desegregation*, Princeton University Press.

Cohodes, S. R. and K. S. Parham (2021): "Charter Schools' Effectiveness, Mechanisms, and Competitive Influence," Working Paper 28477, National Bureau of Economic Research.

Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery (2013): "Barriers to household risk management: Evidence from India," *American Economic Journal: Applied Economics*, 5, 104–35.

Crawford, I. and K. Pendakur (2013): "How many types are there?" *The Economic Journal*, 123, 77–95.

Currie, J. (2006): "The Take Up of Social Benefits," in *Public Policy and the Income Distribution*, ed. by A. J. Auerbach, D. Card, and J. M. Quigley, New York: Russell Sage Foundation, 80–148.

Damera, V. K. (2017): "Choice for the poor or poor choice. India's school choice policy," *Blavatnik School of Government Working Paper, University of Oxford*.

Delavande, A. and B. Zafar (2019): "University Choice: The Role of Expected Earnings, Nonpecuniary Outcomes, and Financial Constraints," *Journal of Political Economy*, 127, 2343–2393.

Deshpande, M. and Y. Li (2019): "Who is screened out? Application costs and the targeting of disability programs," *American Economic Journal: Economic Policy*, 11, 213–48.

Diamond, R., T. McQuade, and F. Qian (2019): "The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco," *American Economic Review*, 109, 3365–3394.

Epple, D., R. E. Romano, and M. Urquiola (2017): "School Vouchers: A Survey of the Economics Literature," *Journal of Economic Literature*, 55, 441–92.

Filmer, D. and L. H. Pritchett (2001): "Estimating wealth effects without expenditure data-or tears: an application to educational enrollments in states of India," *Demography*, 38, 115–132.

Finkelstein, A. and M. J. Notowidigdo (2019): "Take-up and targeting: Experimental evidence from SNAP," *The Quarterly Journal of Economics*, 134, 1505–1556.

Ganimian, A. J., K. Muralidharan, and C. R. Walters (2021): "Augmenting State Capacity for Child Development: Experimental Evidence from India," Working Paper 28780, National Bureau of Economic Research.

Glaeser, E. L. and E. F. P. Luttmer (2003): "The misallocation of housing under rent control," *American Economic Review*, 93, 1027–1046.

Government of India (2011): "Population Census 2011," Tech. rep., Office of the Registrar General & Census Commissioner, India.

Hoxby, C. and S. Turner (2013): "Expanding college opportunities for high-achieving, low income students," *Stanford Institute for Economic Policy Research Discussion Paper*, 12, 7.

Hsieh, C.-T. and M. Urquiola (2006): "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of Public Economics*, 90, 1477–1503.

IIPS (2017): "National Family Health Survey (NFHS-4), 2015–16," Tech. rep., International Institute of Population Sciences.

Imbens, G. W. and D. B. Rubin (1997): "Estimating outcome distributions for compliers in instrumental variables models," *The Review of Economic Studies*, 64, 555–574.

Indus Action (2019): "The Bright Spots Report: Status of social inclusion through RTE Section 12(1)(c)," Tech. rep., Indus Action.

——— (2021): "10 Years of RTE 12(1)(c) in Chhattisgarh," Tech. rep., Indus Action.

Jenkins, S. P., J. Micklewright, and S. V. Schnepf (2008): "Social segregation in secondary schools: how does England compare with other countries?" *Oxford Review of Education*, 34, 21–37.

Jensenius, F. R. (2015): "Development from representation? A study of quotas for the scheduled castes in India," *American Economic Journal: Applied Economics*, 7, 196–220.

Khanna, G. (2020): "Does Affirmative Action Incentivize Schooling? Evidence from India," *The Review of Economics and Statistics*, 102, 219–233.

Kingdon, G. (2020): "The private schooling phenomenon in India: A review," *The Journal of Development Studies*, 1–23.

Kremer, M. (1993): "The O-ring theory of economic development," *The Quarterly Journal of Economics*, 108, 551–575.

Lee, D. S. (2009): "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 76, 1071–1102.

Mbiti, I., K. Muralidharan, M. Romero, Y. Schipper, C. Manda, and R. Rajani (2019): "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania," *The Quarterly Journal of Economics*, 134, 1627–1673.

Mello, U. (2021): "Centralized admissions, affirmative action and access of low-income students to higher education," *American Economic Journal: Economic Policy*.

Mobarak, A. M., P. Dwivedi, R. Bailis, L. Hildemann, and G. Miller (2012): "Low demand for nontraditional cookstove technologies," *Proceedings of the National Academy of Sciences*, 109, 10815–10820.

Monarrez, T. E. (2022): "School attendance boundaries and the segregation of public schools in the US," *American Economic Journal: Applied Economics*.

Muralidharan, K., A. Singh, and A. J. Ganimian (2019): "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 109, 1426–60.

Muralidharan, K. and V. Sundararaman (2015): "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India," *The Quarterly Journal of Economics*, 130, 1011.

Neilson, C., C. Allende, and F. Gallego (2019): "Approximating the Equilibrium Effects of Informed School Choice," Tech. rep., Princeton University, Department of Economics, Industrial Relations Section.

NITI Aayog (2020): "SDG India Index 2019," Report, NITI Aayog.

Otero, S., N. Barahona, and C. Dobbin (2021): "Affirmative Action in Centralized College Admission Systems: Evidence from Brazil," Tech. rep., Working paper.

Pande, R. (2003): "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India," *American Economic Review*, 93, 1132–1151.

Pathak, P. A. (2017): "What really matters in designing school choice mechanisms," *Advances in Economics and Econometrics*, 1, 176–214.

Prakash, N. (2020): "The impact of employment quotas on the economic lives of disadvantaged minorities in India," *Journal of Economic Behavior & Organization*, 180, 494–509.

Pratham (2019): "Annual Status of Education Report," Tech. rep., Pratham.

Rao, G. (2019): "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools," *American Economic Review*, 109, 774–809.

Ravallion, M. (1991): "Reaching the rural poor through public employment: arguments, evidence, and lessons from South Asia," *The World Bank Research Observer*, 6, 153–175.

Reardon, S. F. and A. Owens (2014): "60 years after Brown: Trends and consequences of school segregation," *Annual Review of Sociology*, 40.

Romero, M., J. Sandefur, and W. A. Sandholtz (2020): "Outsourcing Education: Experimental Evidence from Liberia," *American Economic Review*, 110, 364–400.

Singh, A. (2014): "Test score gaps between private and government sector students at school entry age in India," *Oxford Review of Education*, 40, 30–49.

——— (2015): "Private school effects in urban and rural India: Panel estimates at primary and secondary school ages," *Journal of Development Economics*, 113, 16–32.

Söderström, M. and R. Uusitalo (2010): "School Choice and Segregation: Evidence from an Admission Reform," *The Scandinavian Journal of Economics*, 112, 55–76.

Stantcheva, S. (2021): "Understanding Tax Policy: How do People Reason?" *The Quarterly Journal of Economics*, 136, 2309–2369.

Urquiola, M. (2016): "Chapter 4 - Competition Among Schools: Traditional Public and Private Schools," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, Elsevier, vol. 5, 209–237.

Walters, C. R. (2018): "The Demand for Effective charter Schools," *Journal of Political Economy*, 126, 2179–2223.

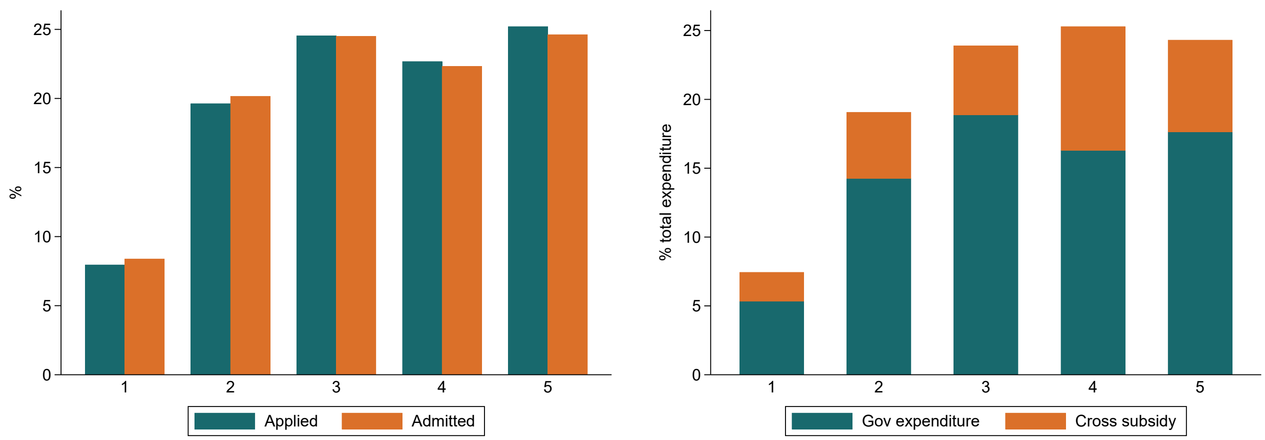
Wiswall, M. and B. Zafar (2018): "Preference for the Workplace, Investment in Human Capital, and Gender," *The Quarterly Journal of Economics*, 133, 457–507.

World Bank (2016): "Chhattisgarh: Poverty, Growth and Inequality," Tech. rep., World Bank.

Figures

Figure 1: Applications, admission, and expenditure by SES quintile

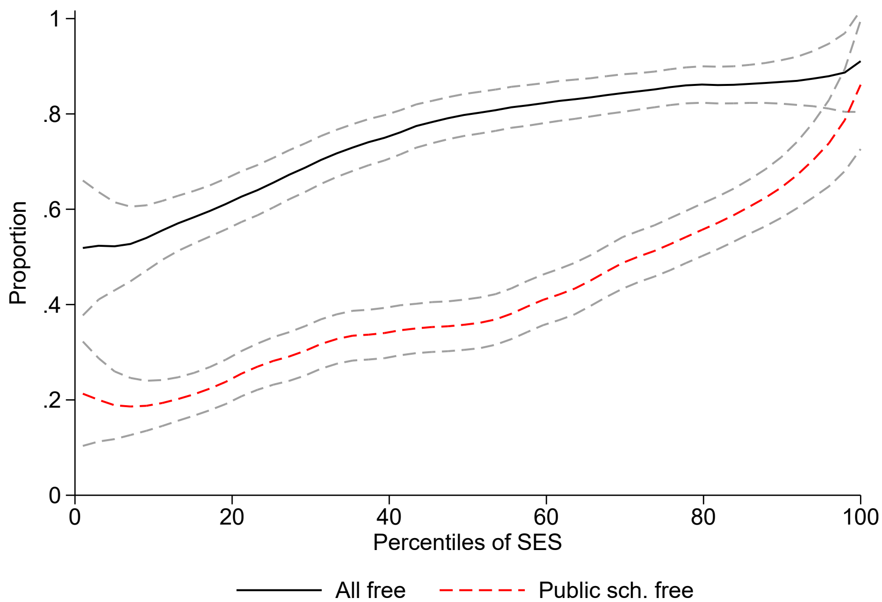
(a) Applications (b) Expenditure



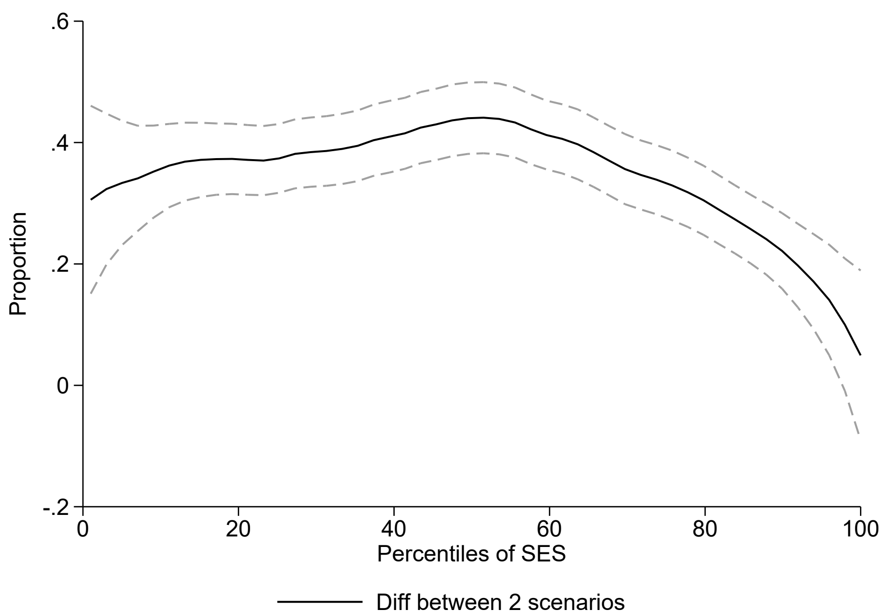
Note: Figure 1a shows the proportion of students that apply and are admitted for an RTE seat that belong to each quintile of the state-wide SES distribution. Quintiles of socioeconomic status are computed based on household ownership of consumer durables and assets which are collected both in the representative National Family and Health Survey 2019-21 for the state-wide population and also in our survey of applicants. Figure 1b shows the proportion of the total costs of RTE quota seats (divided by the reimbursement of the fees given by the government and the cross-subsidy provided by schools if their fees are above the cap on reimbursements) that goes to each quintile of the SES distribution.

Figure 2: Parental demand for private schools in the "voucher" and "status quo" scenarios

(a) Proportion choosing private schools



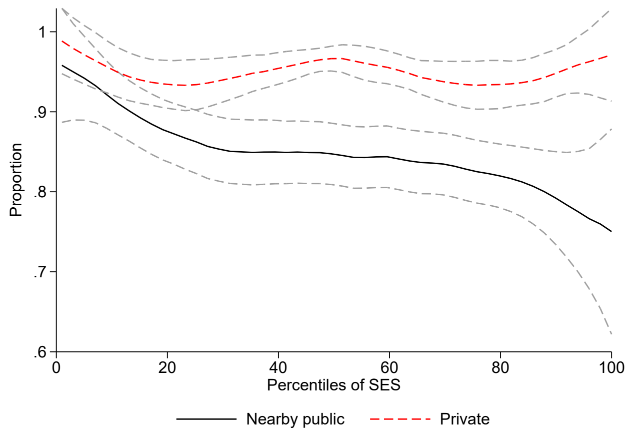
(b) Difference



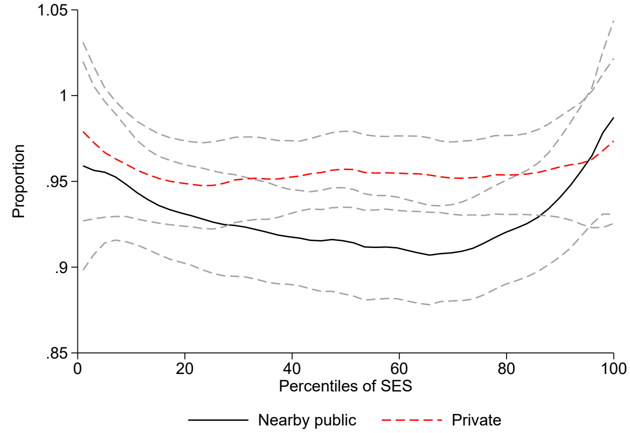
Note: Figure 2a shows local linear regressions which plot, against percentiles of SES, the proportion of parents who choose a private school as the top choice school in (a) "voucher scenario", where all schools are made free for children to attend, and (b) in the "status quo", where public schools are free but private schools charge their posted school fees. Figure 2a shows the difference in proportion of parents choosing private schools in the two scenarios across the SES distribution. We measure socioeconomic status using an index created from household ownership of assets, consumer durables, and quality of housing using Principal Components Analysis (Filmer and Pritchett, 2001). We use an Epanechnikov kernel with a bandwidth of 8 percentiles in all regressions in this plot.

Figure 3: Parental expectations of child experience in private schools

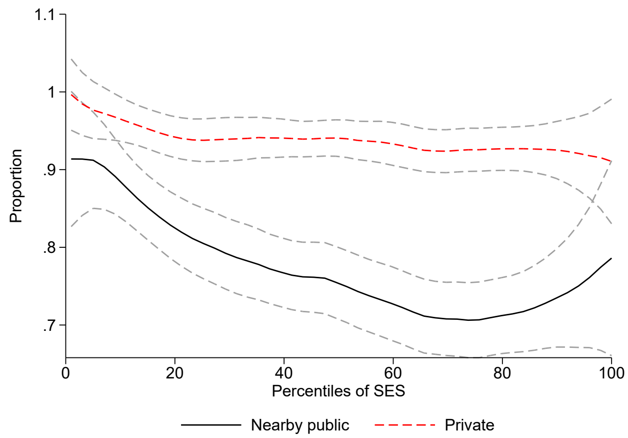
(a) Child likely to be happy



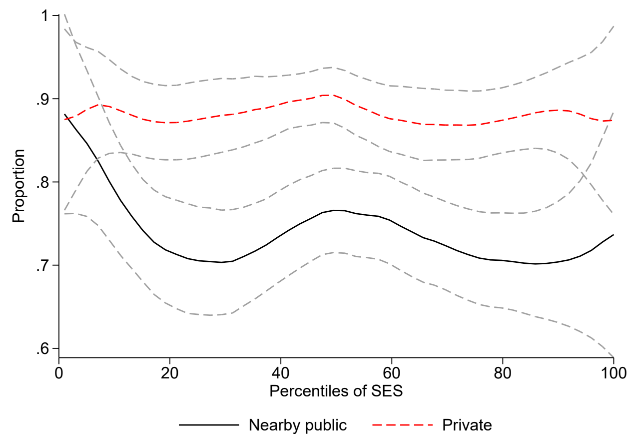
(b) Will have friends, enjoy activities



(c) Teachers will pay attention

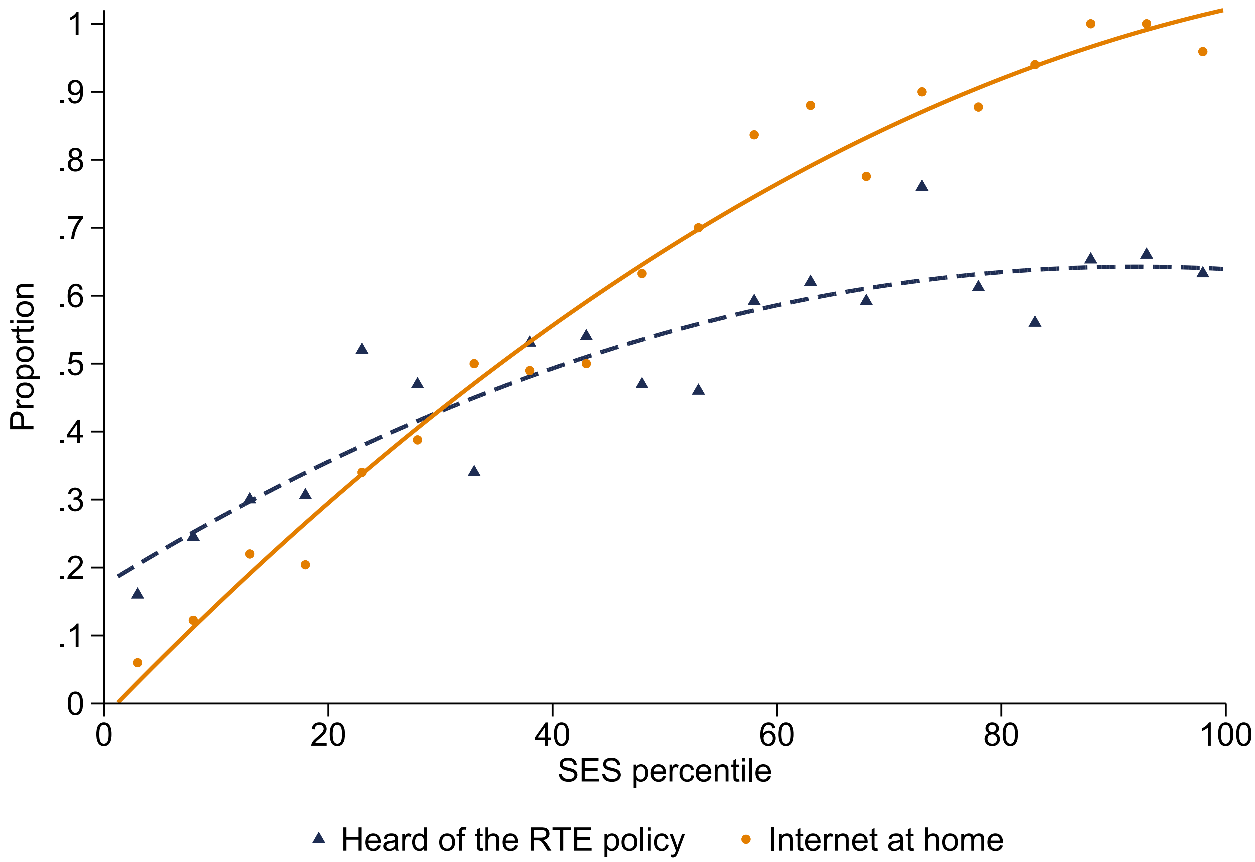


(d) Good job at 30



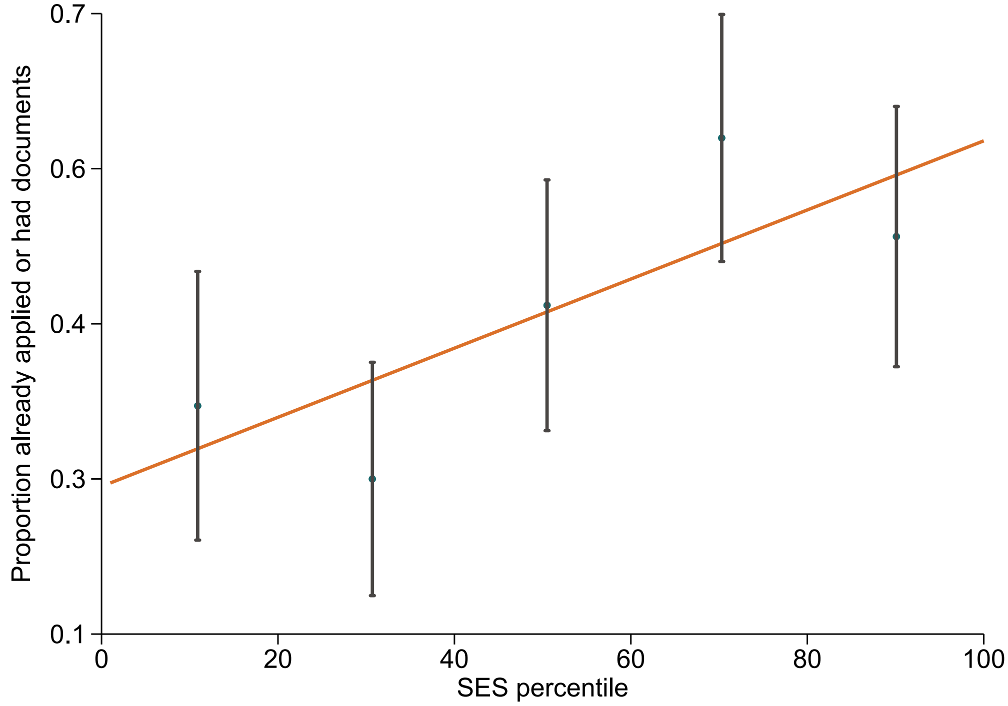
Note: This figure shows output from local linear regression plots relating the percentiles of SES to whether parents responded "Likely" or "Very Likely" to four statements, asked with respect to their child's experience at the nearby government school and a (randomly-chosen) private school: (i) How likely do you think that the child will be happy at his school?; (ii) How likely do you think that the child will have friends and enjoy social activities in school?, (iii) How likely do you think that teachers will pay attention to the child?, (iv) How likely do you think the child will have a good job by the time he is 30?. Parents' responses were elicited on a 5-point scale from "Very Unlikely" to "Very likely" (with additional codes for "Don't know" and "Don't want to answer", which have been removed from the sample for these regressions). All local linear regressions use an Epanechnikov kernel with a bandwidth of 10 percentiles.

Figure 4: Survey evidence on application frictions



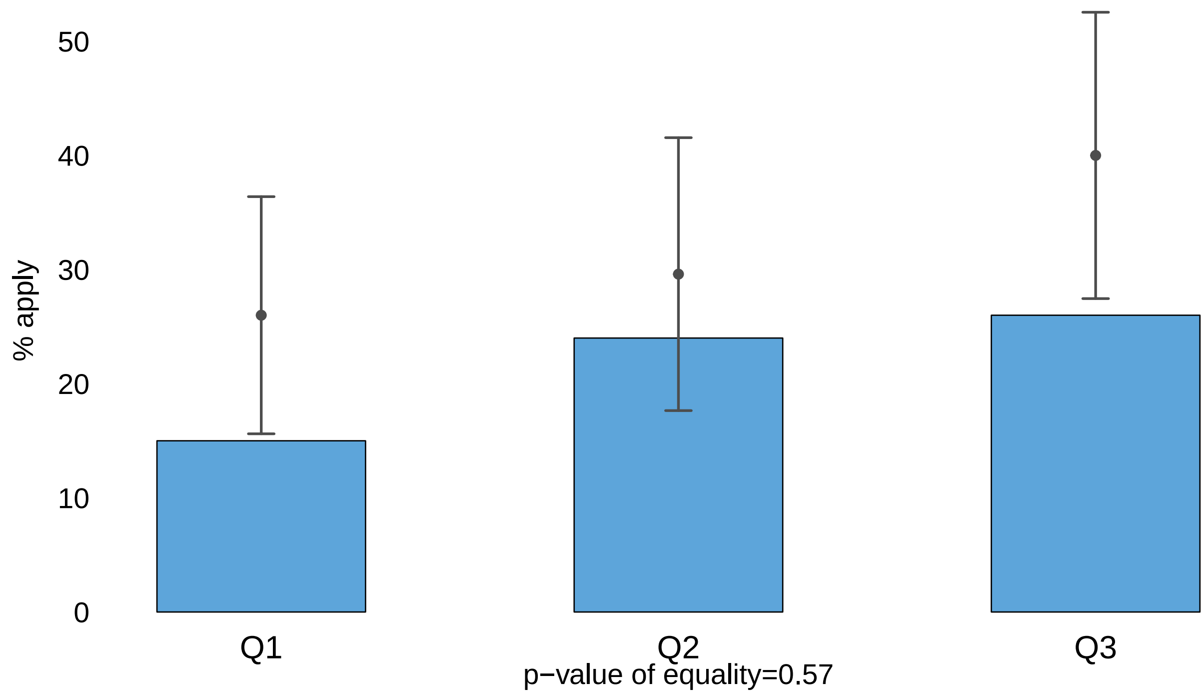
Note: This figure presents binscatter least squares estimations — following Cattaneo et al. (2019) — of the relationship between the percentiles of SES and having heard of the RTE quota policy and having internet at home. We measure socioeconomic status using an index created from household ownership of assets, consumer durables, and quality of housing using Principal Components Analysis (Filmer and Pritchett, 2001).

Figure 5: SES gradient in having all requisite documents for application: Treatment group



Note: This figure plots the proportion of households which had either applied to the RTE quota seats on their own or had all requisite documents, for applying within the randomly-selected treatment group [N=459]. We plot the mean, and associated 95% confidence intervals for each quintile in the socio-economic distribution in our data along with a linear fit.

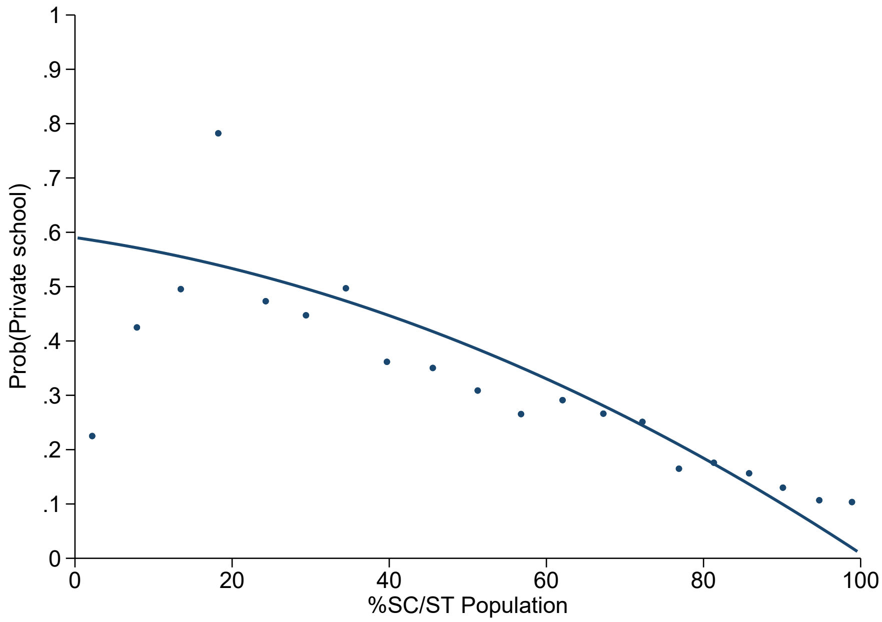
Figure 6: Treatment effects of application support by SES tercile



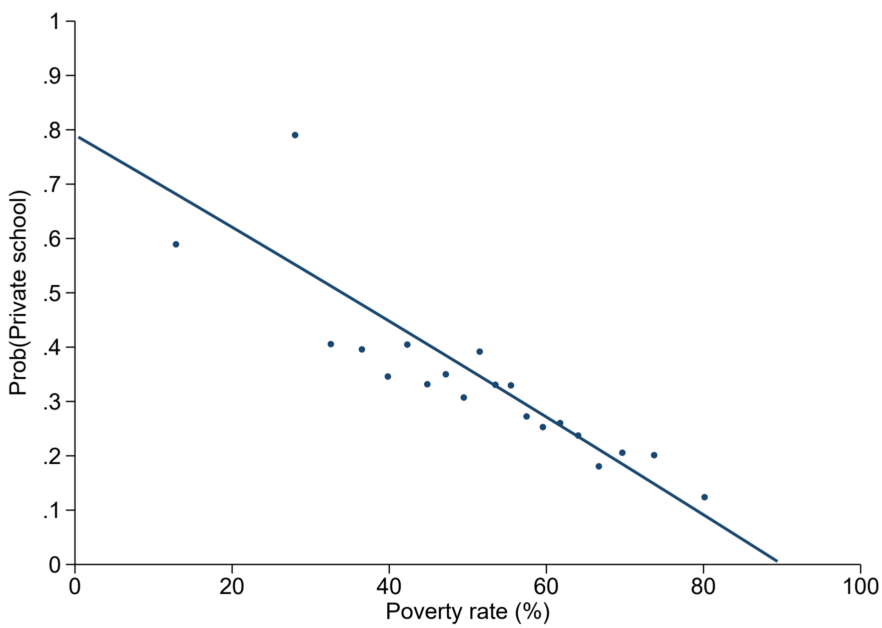
Note: This figure plots the likelihood that households apply for an RTE seat by tercile in the socio-economic distribution. The bars correspond to the control group means in each tercile. Above each bar we present the mean in the treatment group (along with the confidence interval) in each tercile.

Figure 7: Likelihood of a community having a private school by the proportion of eligible population in the community

(a) Scheduled Caste and Schedule Tribe



(b) Poverty rate



Note: Figure 7a presents binscatter least squares estimation — following Cattaneo et al. (2019) — of the relationship between the percentage of the population from Scheduled Castes and Schedule Tribes and the likelihood of having a private school in the community. Figure 7b presents the relationship between the poverty rate and the the likelihood of having a private school in the community. Each observation is a a community (defined by their SHRUG-ID) and is weighted by the total population.

Tables

Table 1: Balance across lottery winners and losers

	Admin data		Phone survey #1		Phone survey #2	
	Control mean	Treatment differential	Control mean	Treatment differential	Control mean	Treatment differential
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.48	0.00	0.49	0.00	0.48	0.01
	(0.50)	(0.01)	(0.50)	(0.02)	(0.50)	(0.02)
	[5,388]	[11,024]	[2,546]	[4,481]	[1,426]	[2,591]
Age (Jan 1st, 2019)	4.06	-0.01	4.00	-0.01	3.98	-0.01
	(0.94)	(0.01)	(0.94)	(0.01)	(0.92)	(0.01)
	[5,388]	[11,024]	[2,546]	[4,481]	[1,426]	[2,591]
Scheduled Caste	0.17	-0.00	0.16	0.01	0.15	0.03*
	(0.38)	(0.01)	(0.37)	(0.01)	(0.36)	(0.01)
	[5,388]	[11,024]	[2,546]	[4,481]	[1,426]	[2,591]
Scheduled Tribe	0.16	-0.00	0.12	-0.00	0.10	-0.00
	(0.37)	(0.01)	(0.32)	(0.01)	(0.30)	(0.01)
	[5,388]	[11,024]	[2,546]	[4,481]	[1,426]	[2,591]
Other Backward Class	0.54	-0.00	0.57	-0.01	0.60	-0.01
	(0.50)	(0.01)	(0.49)	(0.01)	(0.49)	(0.02)
	[5,388]	[11,024]	[2,546]	[4,481]	[1,426]	[2,591]
Rural	0.37	0.00	0.29	0.02**	0.28	0.01
	(0.48)	(0.01)	(0.45)	(0.01)	(0.45)	(0.01)
	[5,388]	[11,024]	[2,546]	[4,481]	[1,426]	[2,591]
Surveyed			0.45	0.02**	0.26	0.03***
			(0.50)	(0.01)	(0.44)	(0.01)
			[5,388]	[11,024]	[5,388]	[11,024]
Allocated a seat			0.22	0.73***	0.18	0.77***
			(0.41)	(0.01)	(0.38)	(0.01)
			[2,487]	[4,446]	[1,305]	[2,675]

Notes: Odd columns report the control (lottery losers) mean, standard deviation of the mean (in parentheses), and number of observations in the control group (in square brackets). Even columns report the treatment effect (difference between lottery winners and losers), the standard error of the effect (in parentheses), and number of observations in the treatment group (in square brackets). Columns 1–2 focus on the full sample. The p-value of the null hypothesis that the differences across all the observable applicant characteristics (Column 2) are jointly zero is .81. Columns 3–4 focus on those who completed the first phone survey. The p-value of the null hypothesis that the differences across all the observable applicant characteristics (Column 4) are jointly zero is .25. Columns 5–6 focus on those who answered our second phone survey. The p-value of the null hypothesis that the differences across all the observable applicant characteristics (Column 6) are jointly zero is .62. All differences control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 2: Effect on the extensive margin of enrollment

	Any school				Private school			
	Control mean	ITT	CCM	LATE	Control mean	ITT	CCM	LATE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All	0.86	0.14***	0.83	0.19***	0.82	0.18***	0.78	0.25***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
		[7,027]		[6,933]		[6,976]		[6,890]
Nursery	0.81	0.19***	0.76	0.26***	0.78	0.22***	0.73	0.29***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
		[3,766]		[3,721]		[3,735]		[3,694]
Kindergarten	0.89	0.11***	0.86	0.16***	0.83	0.17***	0.79	0.23***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
		[1,869]		[1,843]		[1,858]		[1,836]
Grade 1	0.98	0.02***	0.97	0.04***	0.91	0.09***	0.89	0.13***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
		[1,392]		[1,369]		[1,383]		[1,360]

Notes: Columns 1 and 5 report the control (lottery losers) mean and the standard error of the mean (in parentheses). Columns 2 and 6 list the itent-to-treat (ITT) effect (difference between lottery winners and losers), the standard error of the effect (in parentheses), and the number of observations used to estimate the effect (in square brackets). Columns 3 and 7 report the control complier mean (CCM) — the mean outcomes for lottery loser compliers — and the standard error of the CCM (in parentheses). Columns 4 and 8 list the local average treatment effect (LATE) of being assigned an RTE seat (instrumented by winning the lottery), the standard error of the effect (in parentheses), and the number of observations used to estimate the effect (in square brackets). All differences control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 3: Effect on fees

	INR			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Panel A: ITT				
Lottery seat	3,543***	5,037***	2,060***	2,059***
	(226)	(349)	(369)	(419)
Control mean	5,599	5,644	4,782	6,650
Control mean in private	7,704	9,091	6,188	7,624
% out of school (control)	21	34	15	3.1
% in public (control)	6.5	4.2	7.6	9.7
N. of obs.	5,334	2,732	1,556	1,046
Panel B: LATE				
Allocated an RTE seat	4,630***	6,337***	2,764***	2,880***
	(302)	(459)	(501)	(579)
CCM	5,292	5,583	4,386	5,946
CCM in private	7,895	9,766	5,932	7,023
% out of school (CCM)	17	24	14	3
% in public (CCM)	4.8	2.8	6.7	8.1
N. of obs.	5,297	2,719	1,544	1,034

Notes: Fee information comes from administrative data. Students in public schools or not enrolled in school are assigned zero fees. Panel A presents the intent-to-treat (ITT) effect of winning a lottery seat. Panel B presents the local average treatment effects (LATE) of being allocated an RTE (instrumenting with the outcome of the lottery) on the market price of the school a child attends. All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). CCM denotes the mean outcomes for lottery loser compliers. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 4: Effect on the characteristics of the school a child attends

	English medium	% students ST & SC	Facility index	Enrollment	Teachers	PTR
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: ITT						
Lottery seat	.057*	-1.1	.062	48**	.64	2.4*
	(.03)	(1.3)	(.044)	(24)	(.58)	(1.3)
N. of obs.	1,373	980	982	926	954	892
Control mean	0.56	27.73	0.72	397.91	13.24	29.01
Control mean | enrolled	0.57	28.60	0.74	411.54	13.66	30.00
% Enrolled (Control)	98.04	96.96	96.96	96.69	96.90	96.68
Panel B: LATE						
Allocated an RTE seat	.086**	-1.3	.082	63*	.78	3.6*
	(.043)	(1.9)	(.063)	(34)	(.82)	(2)
N. of obs.	1,353	969	971	915	943	881
CCM	0.50	29.60	0.71	421.91	14.21	27.91
CCM | enrolled	0.52	30.50	0.73	438.55	14.75	28.93
% Enrolled (CCM)	97.01	96.65	96.66	96.46	96.47	96.45

Notes: Panel A presents the intent-to-treat (ITT) effects of winning a seat through the lottery on different characteristics of the school the child is enrolled in. The sample is restricted to students applying for seats in Grade 1. Panel B presents the local average treatment effect (LATE) of being allocated an RTE (instrumenting with the outcome of the lottery) on different characteristics of the school the child is enrolled in. CCM denotes the mean outcomes for lottery loser compliers. In Column 1, the outcome is whether the child attends an English medium schools or not. In Column 2, the outcome is the percentage of enrollment taken by Scheduled Castes and Tribes in the school the child attends. In Column 3, the outcome is a principal component analysis (PCA) facility index based on whether the school has computer assisted learning, a homeroom, electricity, a library, a playground, a solid building, a boundary wall, functioning toilets, and solid classrooms. In Columns 4-6 the outcomes are enrollment, number of teachers, and the pupil-teacher ratio (PTR). All columns control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 5: Effect on enrollment in top-choice school

	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Lottery seat at first choice	.56***	.65***	.47***	.43***
	(.012)	(.015)	(.023)	(.03)
N. of obs.	6,293	3,414	1,708	1,171
Control mean	0.39	0.31	0.45	0.50
Control mean | enrolled	0.45	0.39	0.52	0.51
Control mean | enrolled & no RTE seat	0.55	0.53	0.57	0.57
% Enrolled (Control)	85.03	79.98	86.52	97.18
% RTE seat (Control)	29.50	32.97	25.00	25.84

Notes: This table presents the intent-to-treat (ITT) effects of winning a seat in the first-choice school through the lottery on the likelihood of enrolling in this preferred school. All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 6: Government expenditure

	INR			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Market price				
Allocated an RTE seat	4,630***	6,337***	2,764***	2,880***
	(302)	(459)	(501)	(579)
CCM	5,292	5,583	4,386	5,946
CCM in private	7,895	9,766	5,932	7,023
% out of school (CCM)	17	24	14	3
% in public (CCM)	4.8	2.8	6.7	8.1
N. of obs.	5,297	2,719	1,544	1,034
Panel B: Reimbursed fee				
Allocated an RTE seat	6,072***	6,748***	5,078***	5,795***
	(64)	(83)	(111)	(147)
N. of obs.	5,297	2,719	1,544	1,034
Panel C: Non-limit reimbursed fee				
Allocated an RTE seat	9,922***	11,920***	7,151***	8,826***
	(263)	(374)	(479)	(511)
N. of obs.	5,297	2,719	1,544	1,034

Notes: Fee information comes from administrative data. Students in public schools or not enrolled in school are assigned zero fees. Panel A presents the local average treatment effects (LATE) of being allocated an RTE (instrumenting with the outcome of the lottery) on the market price of the school a child attends. Panel B presents the LATE of being allocated an RTE (instrumenting with the outcome of the lottery) on the reimbursed fee (set to zero for children without an RTE seat). Panel C presents the LATE of being allocated an RTE (instrumenting with the outcome of the lottery) on the hypothetical reimbursed fee in the absence of the maximum reimbursement limit (set to zero for children without an RTE seat). All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). CCM denotes the mean outcomes for lottery loser compliers. Table A.10 presents the intent-to-treat (ITT) estimates of winning a lottery seat. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 7: Differences between applicants and average households in Chhattisgarh

	Stochastic Applicants	Deterministic Applicants	All Applicants	NFHS	(2)-(1)	(4)-(3)	(4)-(1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full state sample							
Asset index	0.48	0.25	0.31	0.17	-0.23***	-0.13***	-0.31***
Mother education: above primary	0.81	0.74	0.76	0.58	-0.07***	-0.18***	-0.23***
Father education: above primary	0.87	0.85	0.86	0.71	-0.02	-0.15***	-0.16***
Panel B: Scheduled Caste							
Asset index	0.37	-0.01	0.07	0.17	-0.38***	0.10	-0.20***
Mother education: above primary	0.80	0.78	0.79	0.60	-0.02	-0.18***	-0.20***
Father education: above primary	0.89	0.82	0.83	0.72	-0.07**	-0.11***	-0.17***
Panel C: Scheduled Tribe							
Asset index	0.33	0.26	0.27	-0.28	-0.07	-0.56***	-0.61***
Mother education: above primary	0.72	0.66	0.67	0.41	-0.06	-0.26***	-0.31***
Father education: above primary	0.80	0.90	0.88	0.58	0.10***	-0.30***	-0.23***
Panel D: Other Backward Class							
Asset index	0.55	0.36	0.42	0.41	-0.19***	-0.01	-0.15***
Mother education: above primary	0.83	0.75	0.77	0.67	-0.08***	-0.11***	-0.16***
Father education: above primary	0.88	0.84	0.86	0.77	-0.04**	-0.09***	-0.11***

Notes: This table shows the prevalence of different characteristics for applicant households in our main sample (Column 1), a sample of applicants with no variation in the schools they are assigned to (Column 2), all applicants (a weighted average of Columns 1 and 2, in Column 3), and households in the representative National Family Health Survey (NFHS) 2019-21 sample (Column 4). It also shows the difference between the samples and whether this difference is statistically significant (Columns 5–7). Panel A uses the entire state sample, Panel B focuses on Scheduled Caste households, Panel C focuses on Scheduled Tribe households, and Panel D on Other Backward Caste households. We re-weight our sample to account for differential non-response by household characteristics. We estimate the probability of responding to our survey using a linear probability model that accounts for the household district, caste, and the child's age and gender. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 8: Balance in the application assistance experiment

	Control mean (1)	Treatment differential (2)
Parental edu: Completed primary	0.71	0.02
	(0.46)	(0.03)
	[456]	[464]
Parental edu: No formal education	0.04	0.01
	(0.21)	(0.01)
	[456]	[464]
Improved water source	0.38	0.04
	(0.48)	(0.03)
	[456]	[464]
SES Index	0.09	-0.01
	(0.94)	(0.06)
	[456]	[464]
SC/ST	0.15	0.03
	(0.36)	(0.02)
	[456]	[464]
Private school (when all free)	0.78	-0.02
	(0.42)	(0.03)
	[456]	[464]
Heard about RTE (baseline)	0.52	0.01
	(0.50)	(0.03)
	[456]	[464]
Follow-up interview	0.80	0.00
	(0.40)	(0.03)
	[456]	[464]

Notes: Column 1 presents the control mean, standard deviation of the mean (in parentheses), and the number of observations in the control group (in square brackets). Column 2 reports the treatment effect, the standard error of the effect (in parentheses), and the number of observations in the treatment group (in square brackets). All treatment estimates control for strata (cluster) fixed effects. Standard errors are clustered at the household level. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 9: Treatment effects from application assistance intervention

	(1)	(2)	(3)	(4)
Panel A: Heard of RTE				
Treatment (ITT)	.057***	.057***	.06***	.061***
	(.016)	(.016)	(.016)	(.016)
N. of obs.	739	739	739	739
Control mean	.93	.93	.93	.93
Panel B: Knows application window				
Treatment (ITT)	.18***	.18***	.18***	.18***
	(.029)	(.029)	(.029)	(.029)
N. of obs.	739	739	739	739
Control mean	.69	.69	.69	.69
Panel C: Applied to RTE this year				
Treatment (ITT)	.095***	.096***	.093***	.094***
	(.032)	(.033)	(.033)	(.033)
N. of obs.	739	739	739	739
Control mean	.22	.22	.22	.22
Panel D: Secured RTE seat				
Treatment (ITT)	.033	.032	.032	.033
	(.021)	(.021)	(.021)	(.021)
N. of obs.	739	739	739	739
Control mean	.082	.082	.082	.082
Panel E: Used RTE seat				
Treatment (ITT)	.032	.031	.03	.031
	(.021)	(.021)	(.021)	(.021)
N. of obs.	739	739	739	739
Control mean	.079	.079	.079	.079
SES controls	No	Yes	Yes	Yes
Preferences over private schools	No	No	Yes	Yes
RTE Knowledge (baseline)	No	No	No	Yes

Notes: The outcome in Panel A is whether the household had heard of the RTE policy, in Panel B is whether they knew the right dates for the application window, in Panel C is whether they applied this year, in panel D is whether they secured an RTE seat, and in Panel E is whether they enrolled their children in an RTE seat. Column 1 does not include any additional controls. Column 2 includes socioeconomic status controls (parental education, accessed to improved water access and sanitation, SES index, and caste). Column 3 includes socioeconomic status controls, as well as controls for preferences over private schools (whether children where enrolled in a private school in the past and preferences over private schools in the fictitious scenarios discussed in Section 4.1). Column 4 also controls for knowledge of the RTE policy (whether they had heard of the policy before and whether they had applied for an RTE seat before). All estimations are done via ordinary least squares, controlling for strata (village) fixed effects and clustering standard errors at the household level. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 10: School characteristics across the public and private sector

	Private mean (1)	Public mean (2)	Difference (3)	Difference Block F.E. (4)	Difference Postcode F.E. (5)	Difference Shrug F.E. (6)
% (SC+ST)	35.16	63.12	27.95***	16.45***	16.45***	12.78***
	(40.51)	(32.66)	(0.57)	(0.56)	(0.59)	(0.57)
	[5,377]	[43,476]				
English medium (%)	25.44	0.15	-25.29***	-23.69***	-23.28***	-22.92***
	(43.56)	(3.92)	(0.59)	(0.58)	(0.59)	(0.60)
	[5,377]	[43,476]				
Facility index	0.51	0.02	-0.48***	-0.32***	-0.32***	-0.26***
	(0.69)	(0.84)	(0.01)	(0.01)	(0.01)	(0.01)
	[5,377]	[43,476]				
Enrollment	204.17	69.77	-134.40***	-110.27***	-110.03***	-97.43***
	(232.03)	(56.99)	(3.18)	(2.91)	(2.90)	(3.08)
	[5,377]	[43,476]				
Teachers	11.55	3.84	-7.72***	-6.62***	-6.56***	-6.04***
	(9.38)	(2.25)	(0.13)	(0.11)	(0.11)	(0.12)
	[5,377]	[43,476]				
PTR	19.73	19.35	-0.38	0.61	0.43	0.60
	(27.80)	(16.43)	(0.39)	(0.38)	(0.38)	(0.43)
	[5,377]	[43,476]				
Market share (%)	27	73				

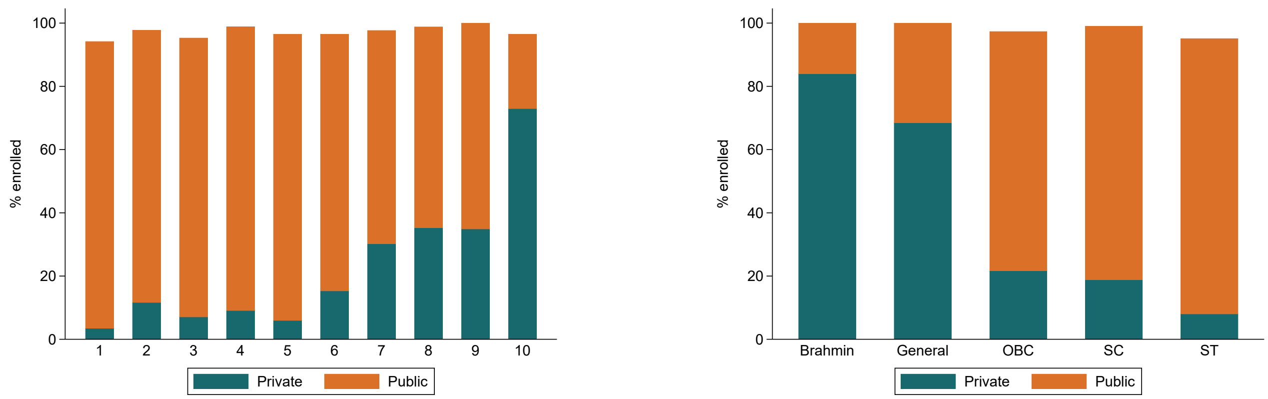
Notes: %(SC+ST) is the percentage of Scheduled Caste or a Schedule Tribe students out of the total enrollment (across all grades). English medium (%) is the percentage of schools with English medium. Facility index is a principal component analysis (PCA) index based on whether the school has computer assisted learning, a homeroom, electricity, a library, a playground, a solid building, a boundary wall, functioning toilets, and solid classrooms. Enrollment is the total size of the school, teachers is the total number of teachers, and PTR is the pupil-teacher ratio. Column 1 shows the mean in private schools (standard deviation in parenthesis, number of observations in square brackets), while Column 2 shows the mean in public schools (standard deviation in parenthesis, number of observations in square brackets). Column 3 presents the difference (with its standard error in parenthesis), Column 4 presents the difference with block fixed effects (with its standard error in parenthesis), Column 5 presents the difference with postal code fixed effects (with its standard error in parenthesis), and Column 6 presents the difference with village/town fixed effects, as defined by the SHRUG-ID created by Asher et al. (2021) (with its standard error in parenthesis). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

For Online Publication

A Additional tables and figures

Figure A.1: Enrollment by household characteristics (2011-12)

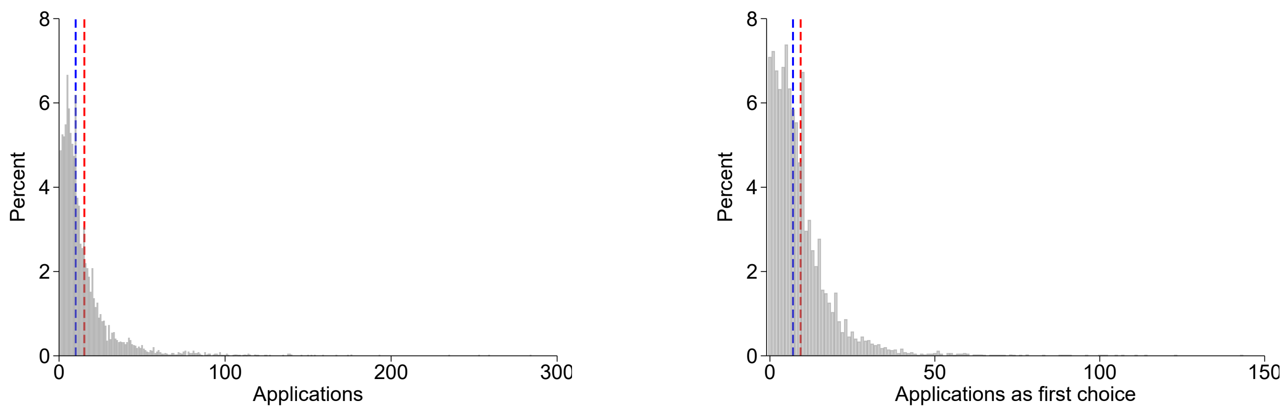
(a) SES Decile (b) Caste



Note: The left panel (Figure A.1a) presents the proportion of children enrolled in different types of schools according to their decile in the income distribution (based on consumption per capita). The right panel (Figure A.1b) presents the presents the proportion of children enrolled in different types of schools according to their caste. Source: Indian Human Development Survey (IHDS)-II (2011-12). See https://doi.org/10.3886/ICPSR36151.v6 *for more details.*

Figure A.2: Applications per school

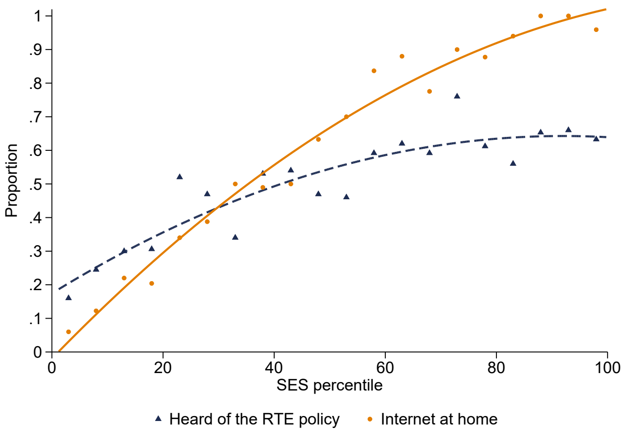
(a) Any ranking (b) Top ranking



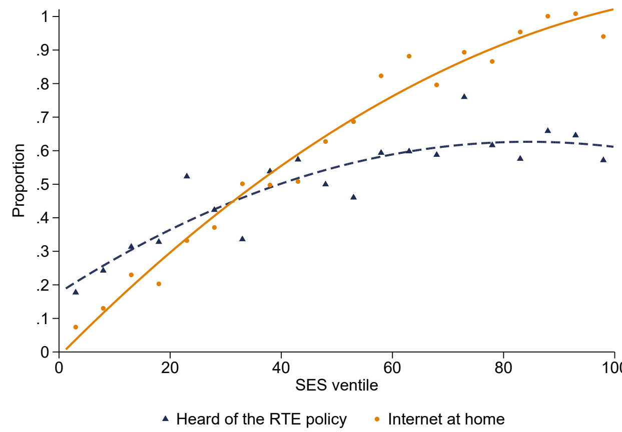
Note: The left panel (Figure A.2a) presents the distribution of the number of times a school is ranked by parents in their preference list. The average (median) school is ranked in 15 (10) applications. The right panel (Figure A.2b) presents the distribution of the number of times a school is ranked first by parents in their preference list. The average (median) school is ranked first in 9.3 (7) applications.

Figure A.3: Socioeconomic status and familiarly with the RTE policy

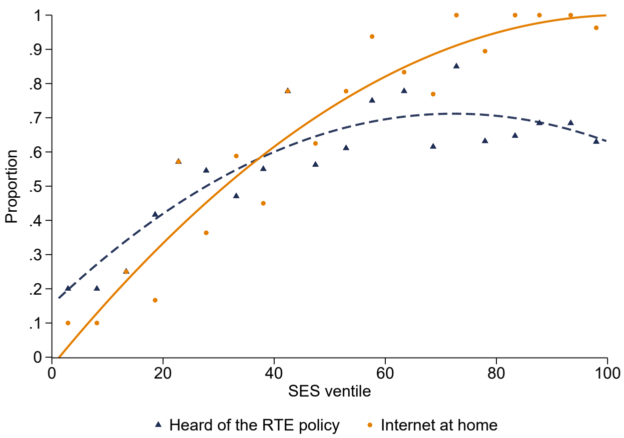
(a) Overall variation



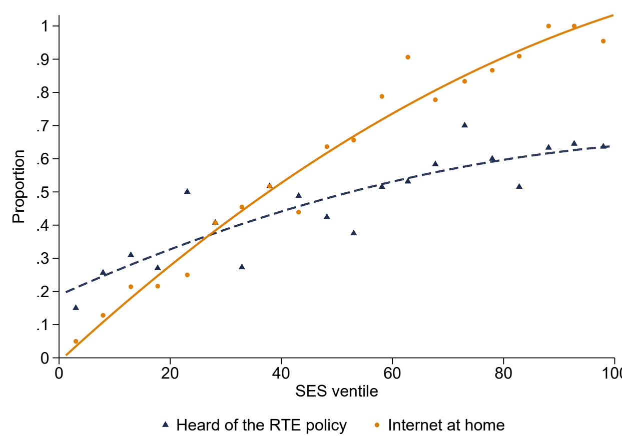
(b) Variation within sampling unit



(c) Urban areas



(d) Rural areas



Note: This figure presents binscatter least squares estimations — following Cattaneo et al. (2019) — of the relationship between the percentiles of SES and: a) having heard of the RTE quota policy and having internet at home. Figure A.3a shows the overall relationship. Figure A.3b shows the relationship within sampling units. Figure A.3c shows the relationship within urban sampling units. Figure A.3d shows the relationship within rural sampling units.

Table A.1: Summary statistics

	Mean	Median	Std. Dev.	Min	Max	N. of obs.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All applicants						
Unalloted	0.12	0.00	0.33	0	1	54,676
Rural	0.56	1.00	0.50	0	1	54,676
Age (Jan 1st, 2019)	4.16	3.97	0.96	2.8	6.3	54,676
Female	0.48	0.00	0.50	0	1	54,676
Scheduled Caste	0.23	0.00	0.42	0	1	54,676
Scheduled Tribe	0.21	0.00	0.40	0	1	54,676
Other Backward Class	0.47	0.00	0.50	0	1	54,676
No. of preferences	1.63	1.00	1.35	1	26	54,676
Panel B: Applicants in phone survey						
Unalloted	0.33	0.00	0.47	0	1	16,703
Rural	0.43	0.00	0.49	0	1	16,703
Age (Jan 1st, 2019)	4.03	3.79	0.94	2.8	6.3	16,703
Female	0.48	0.00	0.50	0	1	16,703
Scheduled Caste	0.18	0.00	0.39	0	1	16,703
Scheduled Tribe	0.18	0.00	0.39	0	1	16,703
Other Backward Class	0.52	1.00	0.50	0	1	16,703
No. of preferences	1.63	1.00	1.21	1	12	16,703
Panel C: Schools						
Seats	10.07	9.14	6.37	1	80	5,863
No. applicants	15.21	10.00	19.01	1	284	5,863
Has nursery seats	0.52	1.00	0.50	0	1	5,863
Has KG seats	0.35	0.00	0.48	0	1	5,863
Has Grade 1 seats	0.47	0.00	0.50	0	1	5,863
Hindi medium	0.52	1.00	0.50	0	1	5,863
English medium	0.41	0.00	0.49	0	1	5,863

Notes: This table presents summary statistics for all lottery applicants (Panel A) and for applicants we attempted to contact during our phone survey (Panel B). Further, we present summary statistics from schools in the lottery (Panel C).

Table A.2: Application behavior by household characteristics

	(1)	(2)	(3)	(4)	(5)
Panel A: Applies to more than one school					
Age (Jan 1st, 2019)	-.02***	-.031***	-.031***	-.03**	-.03**
	(.0032)	(.01)	(.01)	(.013)	(.013)
Female	-.0012	-.0075	-.0078	-.018	-.018
	(.0029)	(.0093)	(.0093)	(.013)	(.013)
Scheduled Caste	-.018**	-.013	-.012	-.038	-.038
	(.0091)	(.023)	(.023)	(.03)	(.03)
Scheduled Tribe	-.022**	.0064	.0091	.012	.012
	(.0093)	(.021)	(.021)	(.03)	(.031)
Other Backward Class	-.015*	.0024	.0033	-.018	-.018
	(.0082)	(.018)	(.018)	(.023)	(.023)
Mother: Education>Primary			.025*	.026	.026
			(.014)	(.019)	(.019)
Asset Index					-.001
					(.0045)
Outcome mean	.29	.39	.39	.41	.41
N. of obs.	53,679	6,596	6,596	3,566	3,566
Panel B: Market price of first choice (ln)					
Age (Jan 1st, 2019)	-.024***	-.043**	-.041**	-.037	-.037
	(.0063)	(.019)	(.019)	(.024)	(.024)
Female	-.0021	-.002	-.0034	-.0038	-.0043
	(.0042)	(.0098)	(.0097)	(.017)	(.017)
Scheduled Caste	-.075***	-.07**	-.067**	-.076**	-.074**
	(.013)	(.029)	(.028)	(.034)	(.034)
Scheduled Tribe	-.1***	-.13***	-.12***	-.12***	-.12***
	(.015)	(.038)	(.037)	(.045)	(.045)
Other Backward Class	-.077***	-.095***	-.092***	-.082***	-.082***
	(.013)	(.028)	(.028)	(.027)	(.027)
Mother: Education>Primary			.08***	.1***	.097***
			(.018)	(.026)	(.026)
Asset Index					.0055
					(.0049)
Outcome mean	8.7	8.9	8.9	9	9
N. of obs.	45,221	5,713	5,713	3,086	3,086
Sample	Admin	Phone 1	Phone 1	Phone 2	Phone 2

Notes: Fee information comes from administrative data. All regressions control for habitation (school cluster households are allowed to apply to) fixed effects. That is, regressions control for the supply of schools available to parents. Panel A has as the outcome whether more than one school was ranked in the application. Panel B contains the market price of the first choice. Column 1 contains the full set of applicants. Columns 2 and 3 restrict the sample to those who answered our first phone survey (when we asked about parental education). Columns 4 and 5 restrict the sample to our second phone survey (when we asked about assets). Standard errors are clustered at the habitation level. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.3: Attrition by child characteristics

	Survey #1	Survey #2
	(1)	(2)
Female	.0034	.0053
	(.0078)	(.0069)
Age (Jan 1st, 2019)	-.013	-.0074
	(.012)	(.01)
Scheduled Caste	-.047***	-.027*
	(.016)	(.014)
Scheduled Tribe	-.12***	-.084***
	(.016)	(.014)
Other Backward Class	-.018	.003
	(.014)	(.012)
Rural	-.066***	-.07***
	(.012)	(.011)
N. of obs.	16,412	16,412
Outcome mean	.44	.26

Notes: The outcome is whether we were able to conduct the interview (=1). All columns control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.4: Compliance

	RTE seat			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Allocated a seat	.76***	.77***	.75***	.73***
	(.01)	(.013)	(.021)	(.025)
N. of obs.	6,959	3,737	1,848	1,374
Control mean	0.18	0.18	0.17	0.19

Notes: This table presents the effect of winning a lottery seat on being allotted an RTE seat. All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.5: Effect on the extensive margin of enrollment, controlling for the probability of being assigned to a private school: Lee bounds and stratas with low attrition

	Strata without attrition		Low attrition strata			Lee bounds	
	ITT	LATE	Differential attrition	ITT	LATE	ITT	
						LB	UB
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: All grades							
Private school (19-20)	0.19***	0.24***	0.02	0.19***	0.26***	0.14	0.29
	(0.05)	(0.07)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
	[214]	[211]	[10,084]	[4,289]	[4,241]	[3,782]	[3,782]
Any school (19-20)	0.14***	0.19***	0.02	0.15***	0.20***	0.10	0.22
	(0.05)	(0.06)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
	[140]	[139]	[10,085]	[4,324]	[4,270]	[3,804]	[3,804]
Panel B: Nursery							
Private school (19-20)	0.26***	0.33***	0.01	0.23***	0.30***	0.16	0.34
	(0.08)	(0.11)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
	[81]	[81]	[4,888]	[2,346]	[2,325]	[2,146]	[2,146]
Any school (19-20)	0.21***	0.28***	0.01	0.20***	0.27***	0.14	0.30
	(0.07)	(0.10)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
	[82]	[81]	[4,888]	[2,368]	[2,344]	[2,160]	[2,160]
Panel C: Kindergarten							
Private school (19-20)	0.13	0.15	0.01	0.17***	0.22***	0.14	0.32
	(0.11)	(0.13)	(0.01)	(0.02)	(0.03)	(0.02)	(0.03)
	[17]	[17]	[3,046]	[1,232]	[1,218]	[961]	[961]
Any school (19-20)	0.13	0.15	0.01	0.12***	0.16***	0.11	0.23
	(0.11)	(0.13)	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)
	[17]	[17]	[3,046]	[1,240]	[1,223]	[966]	[966]
Panel D: Grade 1							
Private school (19-20)	0.07	0.11	0.02	0.11***	0.16***	0.08	0.14
	(0.07)	(0.10)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
	[39]	[39]	[2,151]	[763]	[752]	[675]	[675]
Any school (19-20)	0.02	0.03	0.02	0.04***	0.05***	0.01	0.02
	(0.03)	(0.05)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
	[41]	[41]	[2,151]	[770]	[759]	[678]	[678]

Notes: Columns 1–2 display the results restricting the sample to strata without attrition. Column 1 shows the intention-to-treat (ITT) effect of winning the lottery, and Column 2 the local average treatment effect (LATE) of being assigned an RTE seat (instrumented with winning the lottery). Columns 3–5 report the results after dropping the 25% of the strata with the most differential attrition. Column 3 shows the results of differential attrition, Column 4 the ITT effect, and Column 5 the LATE of being assigned an RTE seat. Columns 6–7 show Lee (2009) style bounds — Column 6 has the lower bound (LB), while Column 7 has the upper bound for (UB) — for the ITT effect of winning the lottery. Standard errors are in parentheses. The number of observations in the treatment effects estimates is in square brackets. All treatment estimates control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.6: Heterogeneity on school enrollment ITT, controlling for the probability of being assigned to a private school

	Any school (19–20)		Private school (19–20)	
	All	Grd 1	All	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Heterogeneity by gender				
Lottery seat	.14***	.029**	.18***	.094***
	(.01)	(.011)	(.012)	(.02)
Female	-.0025	.011	-.00015	.013
	(.014)	(.012)	(.016)	(.024)
Lottery seat × Female	.0035	-.0077	.0014	-.0078
	(.014)	(.012)	(.016)	(.025)
N. of obs.	7,027	1,392	6,976	1,383
Control mean	.87	.98	.82	.92
Panel B: Heterogeneity by parental education				
Lottery seat	.15***	.028***	.19***	.095***
	(.0086)	(.0094)	(.0095)	(.016)
Mother HS	.045**	.018**	.063***	.05**
	(.019)	(.0078)	(.021)	(.025)
Lottery seat × Mother HS	-.05***	-.026***	-.072***	-.058**
	(.019)	(.0098)	(.021)	(.028)
N. of obs.	6,829	1,356	6,784	1,347
Control mean	.87	.98	.82	.92
Panel C: Heterogeneity by caste				
Lottery seat	.15***	.022	.18***	.085***
	(.019)	(.015)	(.021)	(.029)
Other Backward Class (OBC)	.012	.012	.0092	.031
	(.02)	(.014)	(.022)	(.029)
Scheduled Tribe (ST)	-.0053	.0094	-.016	.011
	(.027)	(.015)	(.031)	(.046)
Scheduled Caste (SC)	-.024	-.041	-.049*	-.095**
	(.025)	(.029)	(.029)	(.048)
Lottery seat × OBC	-.016	-.0092	-.011	-.027
	(.02)	(.014)	(.022)	(.03)
Lottery seat × ST	-.0028	-.0053	.014	.0016
	(.027)	(.015)	(.031)	(.046)
Lottery seat × SC	.018	.046	.046	.11**
	(.025)	(.03)	(.029)	(.048)
N. of obs.	7,027	1,392	6,976	1,383
Control mean	.87	.98	.82	.92

Notes: This tables presents the intent-to-treat (ITT) estimates of being assigned a seat by winning the lottery. The outcome in Columns 1–2 is whether the child was enrolled in any school in 2019–2020 (=1). The outcome in Columns 3–4 is whether the child was enrolled in a private school in 2019–2020 (=1). Mother HS indicates whether the mother completed high school. Columns 1 and 3 use the full sample, while Columns 2 and 4 use only Grade 1 students. All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.7: Heterogeneity on school enrollment LATE

	Any school (19-20)		Private school (19-20)	
	All	Grd 1	All	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Heterogeneity by gender				
Allocated an RTE seat	.19***	.044**	.24***	.14***
	(.014)	(.017)	(.016)	(.03)
Female	-.0086	.016	-.0071	.021
	(.018)	(.016)	(.02)	(.034)
Allocated an RTE seat × Female	.0088	-.014	.0072	-.019
	(.019)	(.017)	(.021)	(.036)
N. of obs.	6,933	1,369	6,890	1,360
CCM	.83	.97	.78	.89
Panel B: Heterogeneity by parental education				
Allocated an RTE seat	.2***	.039***	.26***	.13***
	(.012)	(.013)	(.013)	(.023)
Mother HS	.06**	.027**	.084***	.069*
	(.025)	(.011)	(.027)	(.036)
Allocated an RTE seat × Mother HS	-.066**	-.036***	-.095***	-.083**
	(.026)	(.014)	(.029)	(.041)
N. of obs.	6,748	1,335	6,707	1,326
CCM	.83	.97	.78	.89
Panel C: Heterogeneity by caste				
Allocated an RTE seat	.19***	.035	.23***	.13***
	(.024)	(.024)	(.027)	(.045)
Other Backward Class (OBC)	.01	.016	-.00065	.044
	(.025)	(.019)	(.027)	(.04)
Scheduled Caste (SC)	-.041	-.054	-.077**	-.12*
	(.033)	(.04)	(.037)	(.066)
Scheduled Tribe (ST)	-.016	.013	-.036	.011
	(.036)	(.021)	(.041)	(.064)
Allocated an RTE seat × OBC	-.015	-.016	-.0013	-.051
	(.026)	(.021)	(.029)	(.045)
Allocated an RTE seat × SC	.037	.061	.078**	.13*
	(.034)	(.044)	(.038)	(.071)
Allocated an RTE seat × ST	.012	-.01	.039	-.0021
	(.038)	(.024)	(.043)	(.07)
N. of obs.	6,933	1,369	6,890	1,360
CCM	.83	.97	.78	.89

Notes: This table presents the local average treatment effect (LATE) of being assigned an RTE seat (instrumented by winning the lottery). CCM stands for control complier mean — the mean outcomes for lottery losers compliers. The outcomes in Columns 1–2 relate to whether the child was enrolled in any school in 2019–2020 (=1). The outcomes in Columns 3–4 indicate whether the child was enrolled in a private school in 2019–2020 (=1). Mother HS indicates whether the mother completed high school. Columns 1 and 3 use the full sample, while Columns 2 and 4 use only Grade 1 students. All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Table A.6 provides the intent-to-treat (ITT) effect of winning a lottery seat. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.8: Effect of winning different lottery seats on enrollment in the top-choice school

	(1)	(2)	(3)	(4)	(5)	(6)
Won lottery	.46***					
	(.013)					
Won seat in first choice		.56***	.75***	.7***	.77***	.73***
		(.012)	(.017)	(.02)	(.027)	(.03)
Won seat in second choice				-.22***	-.15***	-.19***
				(.026)	(.04)	(.042)
Won seat in third choice						-.24***
						(.043)
N. of obs.	6,293	6,293	2,159	2,159	980	980
Sample	Full	Full	≥ 2 choices	≥ 2 choices	≥ 3 choices	≥ 3 choices

Notes: This table presents the effect of winning different lottery seats on the likelihood of enrolling in the top-choice school. All columns control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.9: Effect on the diversity of the student body

	% SC (1)	% ST (2)	% SC+ST (3)
Panel A: ITT			
Lottery seat	.31	-.15	.15
	(1.7)	(1.5)	(2.3)
Scheduled Tribe	-2.9	19***	16***
	(1.8)	(3.9)	(3.8)
Scheduled Caste	6.9***	-1.9	5.1**
	(2)	(1.7)	(2.5)
Other Backward Class	.41	1.1	1.5
	(1.6)	(1.4)	(2.1)
Lottery seat × Scheduled Tribe	.43	-4.3	-3.9
	(2)	(4.5)	(4.3)
Lottery seat × Scheduled Caste	2.8	-1.3	1.4
	(2.4)	(2.2)	(3)
Lottery seat × Other Backward Class	-.054	-1.7	-1.8
	(1.8)	(2.1)	(2.6)
N. of obs.	980	980	980
Control mean	12.77	14.96	27.73
Control mean | enrolled	13.17	15.43	28.60
% Enrolled (Control)	96.96	96.96	96.96
Panel B: LATE			
Allocated an RTE seat	.51	.34	.86
	(2.5)	(2.3)	(3.4)
Allocated an RTE seat × Scheduled Caste	3.9	-2.2	1.6
	(3.6)	(3.3)	(4.6)
Allocated an RTE seat × Scheduled Tribe	1	-8.4	-7.3
	(3.1)	(6.6)	(6.4)
Allocated an RTE seat × Other Backward Class	.1	-3.1	-3
	(2.6)	(2.9)	(3.8)
Scheduled Caste	5.9**	-.88	5
	(2.9)	(2.4)	(3.6)
Scheduled Tribe	-3.5	23***	19***
	(2.7)	(5.5)	(5.3)
Other Backward Class	.2	2.3	2.5
	(2.2)	(2)	(3)
N. of obs.	969	969	969
CCM	12.25	17.35	29.60
CCM | enrolled	12.64	17.86	30.50
% Enrolled (CCM)	96.65	96.65	96.65

Notes: Panel A presents the intent-to-treat (ITT) effects of winning a seat through the lottery on the proportion of students from Scheduled Castes (SC) and Scheduled Tribes (ST). Panel B presents the local average treatment effect (LATE) of being allocated an RTE (instrumenting with the outcome of the lottery) on the proportion of students from SC and ST. CCM denotes the mean outcomes for lottery loser compliers. All columns control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.10: Intent-to-treat effect on government expenditure

	INR			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Market price				
Lottery seat	3,543***	5,037***	2,060***	2,059***
	(226)	(349)	(369)	(419)
Control mean	5,599	5,644	4,782	6,650
Control mean in private	7,704	9,091	6,188	7,624
% out of school (control)	21	34	15	3.1
% in public (control)	6.5	4.2	7.6	9.7
N. of obs.	5,334	2,732	1,556	1,046
Panel B: Reimbursed fee				
Lottery seat	4,591***	5,295***	3,729***	4,150***
	(77)	(107)	(125)	(187)
N. of obs.	5,334	2,732	1,556	1,046
Panel C: Non-limit reimbursed fee				
Lottery seat	7,529***	9,400***	5,262***	6,315***
	(212)	(301)	(372)	(430)
N. of obs.	5,334	2,732	1,556	1,046

Notes: Fee information comes from administrative data. Students in public schools or not enrolled in school are assigned zero fees. Panel A presents the intent to treat (ITT) effects of being allocated an RTE through the lottery on the market price of the school a child attends. Panel B presents the ITT effects of being allocated an RTE through the lottery on the reimbursed fee (set to zero for children without an RTE seat). Panel C presents the ITT effects of being allocated an RTE through the lottery on the hypothetical reimbursed fee in the absence of the maximum reimbursement limit (set to zero for children without an RTE seat). All regressions control for the probability of being assigned to a private school by the assignment mechanisms following Abdulkadiroğlu et al. (2017). Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.11: Differences in assets between applicants and average households in Chhattisgarh

	Stochastic Applicants	Deterministic Applicants	All Applicants	NFHS	(2)-(1)	(4)-(3)	(4)-(1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full sample							
Owns house	0.88	0.89	0.89	0.64	0.01	-0.25***	-0.24***
number of rooms used for sleeping	2.83	2.76	2.78	2.11	-0.07	-0.66***	-0.72***
table	0.52	0.40	0.43	0.37	-0.11***	-0.06***	-0.14***
cot or bed	0.91	0.91	0.91	0.92	-0.00	0.01	0.01*
chair	0.81	0.77	0.78	0.82	-0.04***	0.04***	0.01
has electricity	0.99	0.99	0.99	0.98	-0.00	-0.01**	-0.01***
electric fan	0.93	0.89	0.90	0.78	-0.03***	-0.12***	-0.14***
has television	0.72	0.58	0.62	0.70	-0.14***	0.08***	-0.03***
has refrigerator	0.25	0.19	0.20	0.22	-0.06***	0.02	-0.03***
has bicycle	0.69	0.78	0.76	0.72	0.10***	-0.04***	0.03***
has motorcycle/scooter	0.52	0.41	0.44	0.59	-0.11***	0.15***	0.07***
Panel B: Scheduled Caste							
Owns house	0.88	0.90	0.90	0.60	0.01	-0.29***	-0.28***
number of rooms used for sleeping	2.67	2.30	2.38	2.03	-0.37***	-0.35***	-0.64***
table	0.48	0.31	0.34	0.33	-0.17***	-0.01	-0.14***
cot or bed	0.90	0.83	0.85	0.94	-0.07**	0.09***	0.04***
chair	0.76	0.74	0.74	0.81	-0.02	0.06**	0.05**
has electricity	0.99	0.98	0.98	0.99	-0.01	0.00	-0.00
electric fan	0.93	0.87	0.88	0.83	-0.06**	-0.06**	-0.11***
has television	0.72	0.53	0.57	0.75	-0.19***	0.18***	0.03
has refrigerator	0.24	0.16	0.18	0.22	-0.08**	0.05*	-0.01
has bicycle	0.70	0.76	0.75	0.68	0.06*	-0.07**	-0.02
has motorcycle/scooter	0.46	0.35	0.37	0.57	-0.10**	0.20***	0.12***
Panel C: Scheduled Tribe							
Owns house	0.93	0.89	0.90	0.63	-0.04	-0.27***	-0.30***
number of rooms used for sleeping	2.96	3.13	3.10	2.05	0.17	-1.05***	-0.91***
table	0.48	0.43	0.44	0.22	-0.05	-0.22***	-0.26***
cot or bed	0.87	0.97	0.95	0.87	0.10***	-0.08***	-0.00
chair	0.80	0.72	0.74	0.71	-0.07	-0.02	-0.08***
has electricity	0.99	0.99	0.99	0.97	0.00	-0.03***	-0.02***
electric fan	0.80	0.80	0.80	0.58	-0.01	-0.22***	-0.22***
has television	0.61	0.41	0.45	0.48	-0.20***	0.03	-0.12***
has refrigerator	0.15	0.21	0.20	0.09	0.06	-0.11***	-0.07***
has bicycle	0.68	0.84	0.81	0.70	0.17***	-0.11***	0.02
has motorcycle/scooter	0.53	0.43	0.45	0.44	-0.10*	-0.01	-0.09***
Panel D: Other Backward Class							
Owns house	0.89	0.89	0.89	0.66	0.00	-0.23***	-0.22***
number of rooms used for sleeping	2.88	2.83	2.84	2.15	-0.05	-0.69***	-0.73***
table	0.53	0.40	0.44	0.45	-0.12***	0.01	-0.08***
cot or bed	0.93	0.92	0.92	0.95	-0.01	0.03***	0.03***
chair	0.82	0.81	0.81	0.88	-0.01	0.07***	0.06***
has electricity	1.00	1.00	1.00	1.00	0.00	-0.00	-0.00
electric fan	0.95	0.94	0.94	0.88	-0.02*	-0.06***	-0.07***
has television	0.76	0.67	0.70	0.80	-0.09***	0.11***	0.05***
has refrigerator	0.26	0.19	0.21	0.27	-0.07***	0.06***	0.01
has bicycle	0.70	0.79	0.76	0.75	0.09***	-0.01	0.05***
has motorcycle/scooter	0.54	0.45	0.48	0.67	-0.09***	0.20***	0.13***

Notes: This table shows the prevalence of different characteristics for applicant households in our main sample (Column 1), a sample of applicants without any variation in the schools they are assigned to (Column 2), all applicants (a weighted average of Columns 1 and 2, in Column 3), and households in the NFHS sample (Column 4). It also displays the difference between the samples and whether this difference is statistically significant (Columns 5–7). Panel A uses the entire sample, Panel B focuses on Scheduled Caste households, Panel C on Scheduled Tribe households, and Panel D on Other Backward Caste households. We re-weight our sample to account for differential non-response by household characteristics. We estimate the probability of responding to our survey using a linear probability model that accounts for the household district, caste, and the child's age and gender. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.12: Differences in maternal education between applicants and average households in Chhattisgarh

	Stochastic Applicants	Deterministic Applicants	All Applicants	NFHS	(2)-(1)	(4)-(3)	(4)-(1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full sample							
no education	0.08	0.10	0.09	0.23	0.02**	0.14***	0.15***
incomplete primary	0.02	0.04	0.04	0.06	0.02***	0.03***	0.04***
complete primary	0.07	0.11	0.10	0.13	0.04***	0.03***	0.06***
incomplete secondary	0.51	0.50	0.50	0.41	-0.01	-0.09***	-0.10***
complete secondary	0.21	0.17	0.18	0.10	-0.04***	-0.08***	-0.11***
higher	0.10	0.08	0.08	0.07	-0.02**	-0.01	-0.02***
Panel B: Scheduled Caste							
no education	0.10	0.06	0.07	0.20	-0.04*	0.13***	0.09***
incomplete primary	0.03	0.04	0.03	0.08	0.01	0.04***	0.05***
complete primary	0.05	0.12	0.10	0.12	0.07***	0.02	0.07***
incomplete secondary	0.50	0.56	0.54	0.44	0.06	-0.11***	-0.06**
complete secondary	0.21	0.15	0.16	0.09	-0.06*	-0.07**	-0.12***
higher	0.10	0.07	0.08	0.07	-0.02	-0.01	-0.03*
Panel C: Scheduled Tribe							
no education	0.11	0.18	0.17	0.36	0.07*	0.20***	0.26***
incomplete primary	0.02	0.03	0.03	0.06	0.01	0.03***	0.04***
complete primary	0.13	0.12	0.12	0.17	-0.01	0.05***	0.04***
incomplete secondary	0.44	0.37	0.39	0.33	-0.07	-0.06***	-0.12***
complete secondary	0.22	0.17	0.18	0.06	-0.05	-0.12***	-0.17***
higher	0.06	0.12	0.11	0.03	0.06*	-0.08***	-0.03***
Panel D: Other Backward Class							
no education	0.06	0.08	0.08	0.16	0.02	0.09***	0.10***
incomplete primary	0.03	0.05	0.04	0.06	0.02**	0.02**	0.04***
complete primary	0.07	0.12	0.11	0.11	0.05***	0.01	0.04***
incomplete secondary	0.53	0.50	0.51	0.46	-0.03	-0.05**	-0.07***
complete secondary	0.21	0.19	0.19	0.13	-0.02	-0.07***	-0.08***
higher	0.09	0.06	0.07	0.08	-0.03***	0.01	-0.01

Notes: This table shows the prevalence of different characteristics for applicant households in our main sample (Columns 1), a sample of applicants without any variation in the schools they are assigned to (Column 2), all applicants (a weighted average of Columns 1 and 2, in Column 3), and households in the NFHS sample (Column 4). It also shows the difference between the samples and whether this difference is statistically significant (Columns 5–7). Panel A uses the entire sample, Panel B focuses on Scheduled Caste households, Panel C on Scheduled Tribe households, and Panel D on Other Backward Caste households. We re-weight our sample to account for differential non-response by household characteristics. We estimate the probability of responding to our survey using a linear probability model that accounts for the household district, caste, and the child's age and gender. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.13: Differences in paternal education between applicants and average households in Chhattisgarh

	Stochastic Applicants	Deterministic Applicants	All Applicants	NFHS	(2)-(1)	(4)-(3)	(4)-(1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Full sample							
no education	0.03	0.03	0.03	0.11	0.00	0.08***	0.08***
incomplete primary	0.02	0.02	0.02	0.07	0.00	0.04***	0.05***
complete primary	0.06	0.09	0.08	0.12	0.03***	0.03***	0.05***
incomplete secondary	0.50	0.50	0.50	0.47	-0.00	-0.02	-0.02**
complete secondary	0.22	0.23	0.22	0.11	0.01	-0.12***	-0.11***
higher	0.15	0.13	0.14	0.12	-0.02**	-0.01	-0.03***
Panel B: Scheduled Caste							
no education	0.04	0.02	0.02	0.11	-0.02*	0.09***	0.07***
incomplete primary	0.03	0.03	0.03	0.07	0.01	0.04***	0.04***
complete primary	0.04	0.13	0.11	0.10	0.09***	-0.01	0.06***
incomplete secondary	0.43	0.39	0.40	0.48	-0.04	0.08**	0.05**
complete secondary	0.26	0.32	0.31	0.12	0.07*	-0.19***	-0.14***
higher	0.21	0.10	0.12	0.13	-0.10***	0.00	-0.08***
Panel C: Scheduled Tribe							
no education	0.07	0.07	0.07	0.20	-0.00	0.13***	0.13***
incomplete primary	0.01	0.02	0.02	0.07	0.01	0.06***	0.06***
complete primary	0.09	0.02	0.03	0.15	-0.08***	0.12***	0.06***
incomplete secondary	0.50	0.54	0.53	0.44	0.04	-0.09***	-0.06***
complete secondary	0.20	0.18	0.19	0.08	-0.01	-0.10***	-0.11***
higher	0.11	0.18	0.17	0.06	0.07*	-0.11***	-0.05***
Panel D: Other Backward Class							
no education	0.02	0.03	0.03	0.06	0.01	0.04***	0.04***
incomplete primary	0.02	0.02	0.02	0.06	-0.00	0.04***	0.04***
complete primary	0.06	0.10	0.09	0.11	0.04***	0.01	0.04***
incomplete secondary	0.52	0.54	0.53	0.51	0.01	-0.03	-0.02
complete secondary	0.22	0.19	0.20	0.12	-0.03*	-0.07***	-0.10***
higher	0.14	0.12	0.13	0.14	-0.02	0.01	-0.00

Notes: This table shows the prevalence of different characteristics for applicant households in our main sample (Columns 1), a sample of applicants without any variation in the schools they are assigned to (Column 2), all applicants (a weighted average of Columns 1 and 2, in Column 3), and households in the NFHS sample (Columns 4). It also shows the difference between the samples and whether this difference is statistically significant (Columns 5–7). Panel A uses the entire sample, Panel B focuses on Scheduled Caste households, Panel C on Scheduled Tribe households, and Panel D on Other Backward Caste households. We re-weight our sample to account for differential non-response by household characteristics. We estimate the probability of responding to our survey using a linear probability model that accounts for the household district, caste, and child's age and gender. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.14: Relationship between familiarity with the RTE quota policy and socioeconomic status

	(1)	(2)	(3)	(4)
Panel A: Heard of the RTE quota seat policy				
SES Index	.14***	.13***	.13***	.14***
	(.012)	(.013)	(.013)	(.014)
Rural		-.1**		
		(.043)		
SC/ST				.08*
				(.04)
N. of obs.	990	990	990	990
Dep. var. mean	.5	.5	.5	.5
Panel B: Internet access				
SES Index	.3***	.29***	.3***	.3***
	(.012)	(.013)	(.013)	(.014)
Rural		-.049		
		(.029)		
SC/ST				.084**
				(.036)
N. of obs.	990	990	990	990
Dep. var. mean	.62	.62	.62	.62
Location fixed effects	No	No	Yes	Yes

Notes: This table shows the relationship between socioeconomic status (measured by a proxy means test), the likelihood that families have heard of the RTE policy (Panel A), and whether families have internet connection at home (Panel B). Column 1 is a simply univariate regression, Column 2 controls for whether the family lives in an urban or a rural location, Column 3 includes sampling location fixed effects, and Column 4 controls for whether the household belongs to one of the eligible caste groups (SC/ST). In rural areas, each sampling location has an area of 100 hectares. In urban areas, each sampling location has 9 hectares. Standard errors are clustered at the sampling location level for all regressions. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.15: Differences between state wide population, applicants, and our in-person survey sample in Raipur

	NFHS All	NFHS Raipur	Applicants from Raipur	In-person sample	(4)-(1)	(4)-(2)	(4)-(3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Summary							
Asset index	0.17	0.59	0.49	0.55	0.38***	-0.04	0.07
Mother education: above primary	0.58	0.64	0.90	0.79	0.21***	0.15***	-0.11***
Father education: above primary	0.71	0.71	0.84	0.78	0.07***	0.07***	-0.06***
Panel B: Assets							
Owns house	0.64	0.73	0.83	0.88	0.24***	0.16***	0.05**
number of rooms used for sleeping	2.11	2.20	2.68	2.79	0.68***	0.59***	0.12
table	0.37	0.53	0.48	0.47	0.10***	-0.06*	-0.01
cot or bed	0.92	0.97	0.93	0.87	-0.06***	-0.11***	-0.06***
chair	0.82	0.87	0.83	0.80	-0.02	-0.07***	-0.03
has electricity	0.98	1.00	1.00	0.98	-0.01	-0.02***	-0.02***
electric fan	0.78	0.94	0.96	0.94	0.16***	0.00	-0.02
has television	0.70	0.89	0.77	0.82	0.12***	-0.07***	0.05*
has refrigerator	0.22	0.35	0.31	0.35	0.13***	-0.01	0.04
has bicycle	0.72	0.77	0.73	0.57	-0.15***	-0.20***	-0.16***
has motorcycle/scooter	0.59	0.74	0.56	0.68	0.09***	-0.06**	0.12***
Panel C: Maternal education							
no education	0.23	0.12	0.01	0.10	-0.12***	-0.02	0.09***
incomplete primary	0.06	0.08	0.06	0.04	-0.02***	-0.04***	-0.02
complete primary	0.13	0.15	0.02	0.07	-0.06***	-0.09***	0.04***
incomplete secondary	0.41	0.38	0.64	0.47	0.06***	0.09***	-0.17***
complete secondary	0.10	0.10	0.20	0.19	0.09***	0.08***	-0.01
higher	0.07	0.16	0.06	0.14	0.06***	-0.02	0.07***
Panel D: Paternal education							
no education	0.11	0.06	0.01	0.06	-0.05***	0.00	0.06***
incomplete primary	0.07	0.11	0.01	0.04	-0.02***	-0.07***	0.03***
complete primary	0.12	0.12	0.13	0.11	-0.01	-0.01	-0.03
incomplete secondary	0.47	0.39	0.55	0.46	-0.01	0.08**	-0.09***
complete secondary	0.11	0.11	0.22	0.16	0.05***	0.05**	-0.06**
higher	0.12	0.22	0.07	0.16	0.04***	-0.06**	0.09***

Notes: This table shows the prevalence of different characteristics for households in NFHS sample (Columns 1), households in Raipur in the NFHS sample (Column 2), all applicants from Raipur, and the sample of households in Raipur we interviewed in person (Columns 4). It also shows the difference between the samples and whether this difference is statistically significant (Columns 5–7). We re-weight our sample to account for differential non-response by household characteristics. We estimate the probability of responding to our (phone) survey for applicants using a linear probability model that accounts for the household district, caste, and child's age and gender. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table A.16: Heterogeneity of treatment effects from application assistance intervention by household socioeconomic status

	Applied (1)	Secured seat (2)
Panel A: Above/below median		
Treatment (ITT) × SES - Below median	.086*	.019
	(.046)	(.029)
Treatment (ITT) × SES - Above median	.12**	.051
	(.051)	(.034)
N. of obs.	689	689
p-value (H_0: all equal)	.63	.47
Panel B: Terciles		
Treatment (ITT) × SES - Tercile 1	.11**	.036
	(.053)	(.036)
Treatment (ITT) × SES - Tercile 2	.056	.005
	(.061)	(.039)
Treatment (ITT) × SES - Tercile 3	.14**	.068
	(.064)	(.044)
N. of obs.	689	689
p-value (H_0: all equal)	.64	.57
p-value(H_0:Extremes are equal)	.72	.58
Panel C: Quintiles		
Treatment (ITT) × SES - Quintile 1	.13**	.047
	(.067)	(.039)
Treatment (ITT) × SES - Quintile 2	.052	-.012
	(.074)	(.052)
Treatment (ITT) × SES - Quintile 3	.018	-.0059
	(.079)	(.049)
Treatment (ITT) × SES - Quintile 4	.28***	.19***
	(.082)	(.062)
Treatment (ITT) × SES - Quintile 5	.044	-.035
	(.082)	(.051)
N. of obs.	689	689
p-value (H_0: all equal)	.13	.033
p-value(H_0:Extremes are equal)	.39	.2

Notes: The outcome in Column 1 is whether the household applied this year and in Column 2 is whether they secured an RTE seat. All regressions include socioeconomic status controls, as well as controls for preferences over private schools (whether children were enrolled in a private school in the past and preferences over private schools in the fictitious scenarios discussed in Section 4.1), as well as controls for knowledge of the RTE policy (whether they had heard of the policy before and whether they had applied for an RTE seat before). All estimations are done via ordinary least squares, controlling for strata (village) fixed effects and clustering standard errors at the household level. We exclude observations without all the necessary data to generate the SES index. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

B Control complier mean

We will compare the LATE estimate to the control compliers mean — the mean outcomes for compliers who lose the lottery (and therefore do not get an RTE seat through other means). This is the relevant comparison, as it is the counterfactual outcome for compliers (over which the LATE is estimated). To do so, we follow Imbens and Rubin (1997) and Abadie (2003) (and specifically Abdulkadiroğlu et al. (2018)'s implementation of Lemma 2.1 in Abadie (2002)). Intuitively, the mean outcome for those without an RTE seat is a weighted combination of the mean outcome for never-takers and for compliers who lost the lottery; the weights correspond to the share of these subpopulations in the entire population, which we can infer from the data. Since we can also infer the mean outcome for never-takers by studying those who won the lottery but do not have an RTE seat, we can back out the mean outcome for compliers who lost the lottery.

Specifically, let $Y_i(1)$ and $Y_i(0)$ denote the potential outcome for individual i as a function of whether they were allotted an RTE seat. Let $T_i(1)$ and $T_i(0)$ denote the potential treatment (being allotted an RTE seat), as a function of the outcome of the lottery (Z_i). The mean value of $g(Y_i)$ for compliers who lose the lottery is:

$$\mathbb{E}[g(Y_i(0))|T_i(1) > T_i(0)] = \frac{\mathbb{E}[g(Y_i)(1-T_i)|Z_i = 1] - \mathbb{E}[g(Y_i)(1-T_i)|Z_i = 0]}{\mathbb{E}[1-T_i|Z_i = 1] - \mathbb{E}[1-T_i|Z_i = 0]} \tag{4}$$

Setting $g(Y_i) = Y_i$ we obtain the average control outcome for compliers (i.e., $\mathbb{E}[Y_i(0)|T_i(1) > T_i(0)]$). This quantity can be estimated via two-stage least-squares by regressing the interaction of the outcome (Y_i) with an indicator for not being assigned an RTE seat ($1 - T_i$) on an indicator for not being assigned an RTE seat, using the outcome of the lottery as an instrument.

C Survey measures of parental demand and expectations

In 2022, we surveyed a representative sample of households in Raipur district to elicit information about parental preferences and expectations about different types of schools and their information about the RTE quota policy (described in Section 2). Here, we provide some additional detail about these elicitation procedures.

C.1 Eliciting parental preferences

The central part of the data collection is a hypothetical choice exercise, modeled after Delavande and Zafar (2019) but adapted to primary school markets in Raipur. The list of characteristics, and the value assigned to each type of school, was taken from administrative data.

C.1.1 Schooling options

We presented characteristics for five schools to each household:

1. **Nearby public school:** Enrollment, number of teachers/classrooms, and highest grade were fixed near the 25th percentile of school enrollment in the public sector.

2. **Distant public school:** Enrollment, number of teachers/classrooms and highest grade were fixed near the 75th percentile of school enrollment in the public sector.

3. **Lower-range private school:** Enrollment, number of teachers/classrooms and highest grade in the school benchmarked to the bottom tercile of private school fees.

4. **Mid-range private school:** Enrollment, number of teachers/classrooms and highest grade in the school benchmarked to the middle tercile of private school fees.

5. **Expensive private school:** Enrollment, number of teachers/classrooms and highest grade in the school benchmarked to the top tercile of private school fees.

We randomized the distance of each school to the household: for School 1, this was randomized to be between 250-750 meters in increments of 250m; for all other schools, this was randomized to be between 500-2000 meters, in increments of 500 meters. Reflecting actual prevalence, the medium of instruction was fixed as Hindi for government schools and English for Schools 4 and 5, and randomized between them for School 3. The precise school fees shown to households for each of the schools was chosen randomly from a range of values. This was created using the p10-p90 spread in each tercile but adding and subtracting INR 1,000 from the boundaries to have some overlap in fee distributions across schools.

The specific values for each characteristic were computed separately for rural and urban areas in Raipur and are presented in Table C.1. These characteristics were presented to households using a visual stimulus (see Figure C.1) and a surveyor explained each of the characteristics in detail before eliciting responses.

Table C.1: School characteristics shown to respondents

Characteristic Sector	School 1 Public	School 2 Public	School 3 Private	School 4 Private	School 5 Private
			URBAN AREAS		
Enrollment	75	200	190	150	230
Num. of teachers	3	9	12	12	17
Num. of classrooms	5	8	8	8	9
Highest grade offered	5	5	8	8	12
Medium of instruction	Hindi	Hindi	Hindi	Hindi/English	English
Annual fees (in Rupees)	0	0	2000-5500	3500-8500	6500-50000
Distance	250-750 m	500-2000 m	500-2000 m	500 - 2000 m	500-2000 m
			RURAL AREAS		
Enrollment	75	150	170	140	120
Num. of teachers	3	7	11	12	11
Num. of classrooms	4	6	8	8	8
Highest grade offered	5	5	8	8	8
Medium of instruction	Hindi	Hindi	Hindi	Hindi/English	English
Annual fees (in Rupees)	0	0	2000-4500	2500-8000	6000-20000
Distance	250-750 m	500-2000 m	500-2000 m	500 - 2000 m	500-2000 m

Notes: This table shows the specific values for each characteristic shown to households in rural and urban areas. Some survey characteristics were randomized within realistic ranges observed in the data. The medium of instruction was randomized for School 3 (cheaper private schools) between Hindi and English. Fees were randomized in a range within each private option, as shown above. Distance was randomized in increments of 250 meters for School 1 and 500 meters for the other 4 schools.

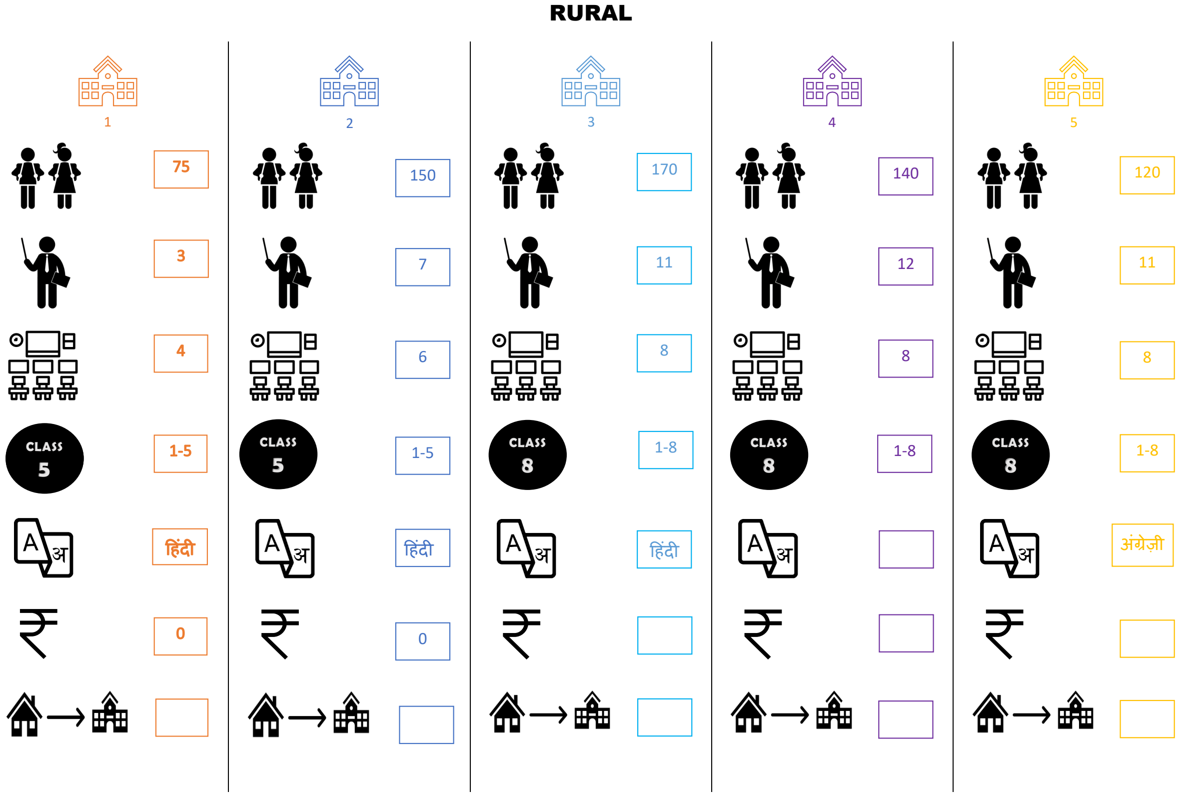
Parents were asked to rank these schools under three different scenarios[57]:

- **Status quo:** Public schools are free to attend, private schools charge their sticker price
- **RTE policy:** One private school (randomly selected) is made free to attend, the others charge their sticker price
- **Unconditional voucher:** All 5 schools are free to attend for the household.

We envisaged the comparison between the "status quo" and the "unconditional voucher" scenario as a measurement of financial constraints. The RTE scenario resembles the actual design of the policy and we use it to study substitution across schools when only one is made free.

[57]Where schools have a positive price, the households were also allowed to answer that they could not afford to send their child to the school. This option was added after piloting, when it became apparent that parents wanted to distinguish this from rating a school as bottom-ranked. This distinction does not matter for our purposes since we only focus on the top-ranked choice in each scenario in our analysis.

Figure C.1: Example of stimulus shown to parents for stated choice exercise



Note: This figure shows the response stimulus as shown to individuals. Fields that are empty correspond to values that were generated randomly. In addition to information shown above, surveyors informed respondents that Schools 1 and 2 were government schools, while Schools 3-5 were private schools.

C.2 Validating survey data on stated choices

We piloted the instrument extensively to ensure respondent comprehension. Results in the final sample also have considerable face validity: Most of the sample chooses private schools in the unconstrained voucher scenario than in status quo (where private schools have a positive price), with financial constraints binding more for individuals at lower SES percentiles (see Figure 2).

We carry out further validation by examining the proportion of households whose stated choices across scenarios violate the General Axiom of Revealed Preference (GARP). While it is common to have *some* respondents violating GARP in most datasets (Crawford and Pendakur, 2013), a large share of violators in this simple discrete choice case, where constraints are made clear, would suggest respondent incomprehension and a general lack of construct validity.

We test for two types of GARP violations[58]:

[58]This exercise implicitly assumes that receiving a scholarship does not enter the utility function of

1. Respondents choose any private school in status quo, but choose a public school when all schools are free
2. Respondents rank a specific private school higher in status quo than when it is made free in the RTE quota scenario (but all other schools have the same price as before)

Overall, we find that stated choices of the vast majority of households respect GARP restrictions. Only 2.3% of households (N=25) choose a private school in status quo but a public school in the voucher scheme; this contrasts with 37.7% of the sample (N=400) who switch from public schools in status quo to private schools in the voucher scenario, or the 59.8% (N=634) who stay in the same sector in both scenarios.

Results are also reassuring for the second type of potential violation. Only 3.4% of respondents (N=37) rate a school lower when offered for free in the RTE quota scenario than in the status quo, compared to 62% (N=659) who rate it higher (with the remainder leaving the ranking unchanged across scenarios). The magnitude of such violations declines with the size of the price discount. The proportion of violators is 2.2% for the sample offered the most expensive private school, 4% for the mid-price school and 4.5% for the low-price private school. Thus respondent comprehension is high across scenarios, although there is suggestive evidence that their attention responds to (hypothetical) stakes in the choices.

Overall, we find high internal consistency in individuals' reported choices across scenarios. Combined with a sensible relationship between school choice and socioeconomic status, comparing individuals both within and across scenarios, this indicates that our survey captured meaningful variation in household preferences and constraints.

C.3 Types of take-up of RTE seats

In the main paper, we only use the survey data on stated choices to understand absolute levels of private school demand and financial constraints across the socioeconomic distribution. Our data however allow for generating richer insights into take-up.

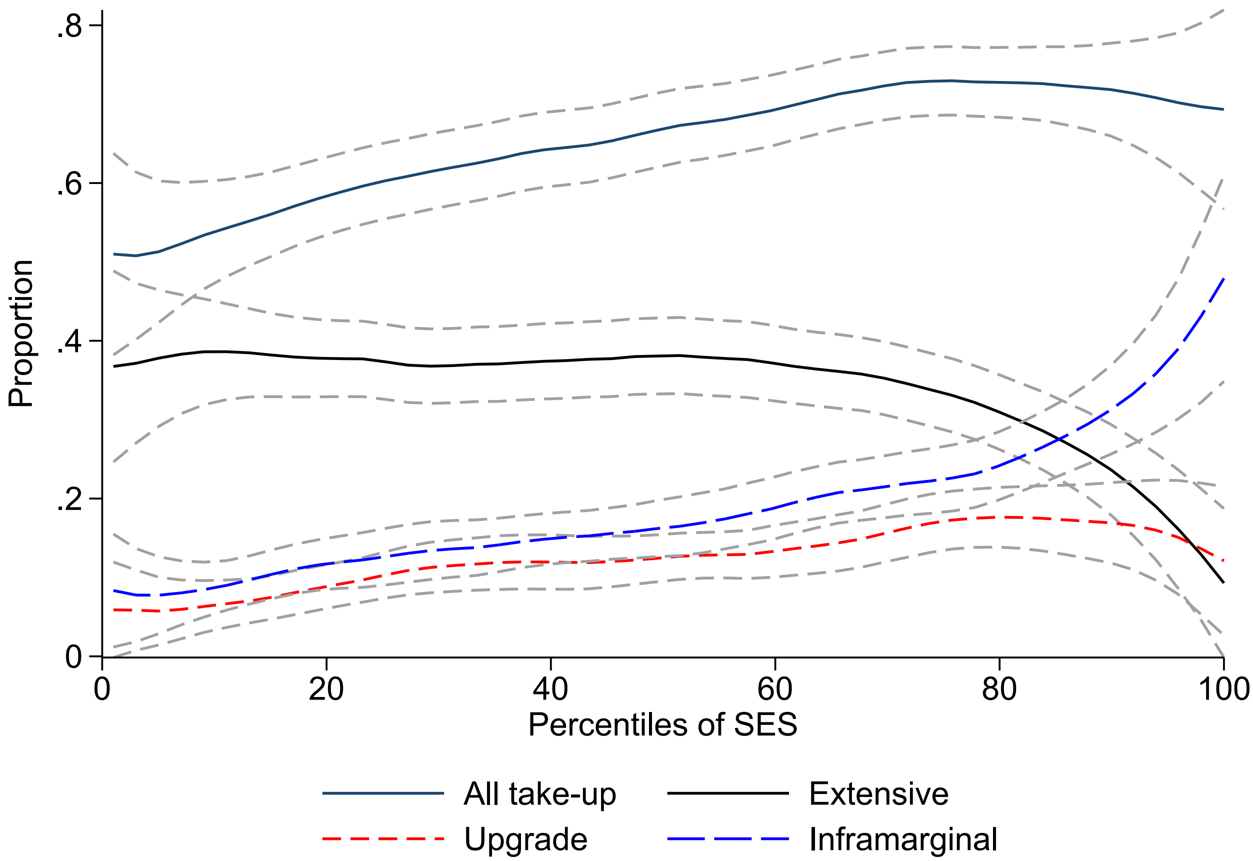
Comparing the offer of a free seat at *one* (randomly-chosen) private school in the "RTE quota" scenario, which reflects policy design more closely, to the "status quo" scenario provides further insights into take-up. We define take-up as choosing, in the RTE quota scenario, the school offered for free as the top choice. Specifically, it provides an initial estimate of what proportion of households across the SES distribution would accept RTE quota seat offers and whether they would use RTE quota seats to move into the private sector ("extensive margin shift"), upgrade within the private sector to a more expensive school, or be fully inframarginal (i.e., attend the same or a less-expensive school).

Figure C.2 plots the proportion of the sample across the SES distribution that corresponds to each type of take-up. Take-up is 52% in the bottom quintile, rises linearly with SES until about the 75th percentile of SES, and flattens out thereafter at about 73%. At low-SES levels, take-up below 100% reflects households that prefer public schools even in the unconstrained "voucher" scenario; at the top end, it reflects that more

households (i.e., that households would not prefer to pay for a good rather than receive it for free). This restriction is reasonable in this setting but can be violated if there is substantial stigma attached to scholarships.

parents prefer and can afford to pay for more expensive private schools than the randomly-offered private school. Reflecting counterfactual status quo choices, nearly all take-up in the bottom quintile consists of extensive margin shifts while, at the top of the socioeconomic distribution, a large share is fully inframarginal.[59]

Figure C.2: Types of take-up of a free seat in "RTE quota" scenario



Note: This figure plots local linear regression plots which relate the percentiles of SES to four quantities in the "RTE quota" scenario (where one, randomly-chosen private school is made free but others charge tuition fees at posted rates): (i) households who choose the free private school as their top choice ("All take-up"), (ii) households which choose a public school in "status quo" but the free private school in RTE quota scenario ("Extensive margin"), (iii) households which choose a cheaper private school in "status quo" but the free school in the "RTE quota" scenario ("Upgrade") and (iv) households which choose the same or a more expensive private school in "status quo" but the free school in the "RTE quota" scenario ("Fully inframarginal"). We measure socioeconomic status using an index created from household ownership of assets, consumer durables, and quality of housing using Principal Components Analysis (Filmer and Pritchett, 2001). All local linear regressions use an Epanechnikov kernel with a bandwidth of 10 percentiles.

[59]A substantial share of take-up in the top quintile reflects *downgrades* — choosing to attend a school that was lower-ranked than other private schools in status quo but is now most-preferred when offered for free.

C.4 Parental expectations about school-specific experience and outcomes

We were concerned, ex-ante, that poorer households may differ substantially in their expectations about the experience that their children might have in private schools. This could, in principle, be a reason for their low application rates.

Our approach to survey measurement of these expectations was through direct questions. We elicited expectations on the following four dimensions, each of which was collected using a 5-point Likert scale going from "Very Unlikely" to "Very likely":

- How likely do you think that the child will be happy at his school?
- How likely do you think that the child will have friends and enjoy social activities in school?
- How likely do you think that teachers will pay attention to the child?
- How likely do you think the child will have a good job by the time he is 30?

We elicited these responses for only two schools — the nearby government school and one (randomly chosen) private school out of the three private options. This was done to limit survey burden for respondents.

In addition to the 5-point Likert scale responses, respondents could also answer "Can't Say' and "Don't Want to Answer". We treat these last two responses as missing data. This is not a major problem for the first three questions where but is a substantial issue for the last question (about future job prospects) where about 30% of households use the "Can't Say" option. Given the substantial difference in the usage of this option across the four questions, we think this reveals genuine uncertainty for households in forecasting labor market outcomes for children who are currently of preschool age. The proportion of missing responses is very similar in each question across the public and private options. Thus, this pattern does not pose a problem in our current analysis (which focuses on a descriptive comparison of expectation across sectors).[60]

[60] However, adapting these survey questions for use in models where a precise value of the expectation is required to rationalize household choices may need more survey adaptation.

D Simulating potential reallocation of RTE quota seats

We conduct two simulation exercises to understand the extent to which the RTE quota policy may change the caste composition across the public and private sectors.

Taking the status quo as the benchmark, we compute school composition in two alternative scenarios to obtain a rough estimate of how effects on school integration depend on undersubscription and reallocation (see Table D.1). The first scenario assumes that all currently used seats are being used by SC/ST children and reallocates these children to public schools in the same "neighborhood". We define "neighborhood" in two different ways. The results are qualitatively similar either way. First, we use postal codes to define a neighborhood. Postal codes come from the U-DISE dataset, but in India postal codes are larger than in many other settings. The second alternative maps school coordinates to SHRUG IDs (Asher et al., 2021) and defines each ID as a different neighborhood. This first reallocation exercise approximates the worst-case scenario for the private-public difference in caste composition of schools if the policy had zero take-up.[61]

In the second scenario, we take every currently unfilled quota seat in each private school and move an SC/ST student from a public school in the neighborhood: this approximates the best-case scenario of addressing demand-side frictions and ending undersubscription.

In the status quo, the share of SC/ST students in the public sector is $\sim$28 percentage points higher than in the government sector. In Scenario 1, this increases to 34 percentage points, while in Scenario 2, it reduces to 15 percentage points. Further, the best-case scenario reduces the private-public gap within postcodes close to zero. Overall, although spatial constraints limit the equalization of caste composition in even the best-case scenario, it appears that substantial reductions in segregation remain possible.

[61] This exercise is distinct from imagining a situation where the policy itself did not exist since we take the capacity, price and characteristics of schools as given. The introduction of a large policy of this nature likely had market-level effects on these variables, as well as private school entry and exit.

Table D.1: Proportion of Scheduled Caste and Scheduled Tribe students in public and private schools under different scenarios

	Private mean (1)	Public mean (2)	Difference (3)	Difference Block F.E. (4)	Difference Postcode F.E. (5)	Difference Shrug F.E. (6)
Panel A: Movements within postal code						
% (SC+ST)	28.21	56.37	28.17***	20.02***	19.96***	17.03***
	(21.69)	(33.13)	(0.48)	(0.40)	(0.41)	(0.47)
	[5,183]	[30,221]				
% (SC+ST-Used Seats)	23.15	57.24	34.10***	26.44***	26.60***	24.02***
	(22.99)	(32.87)	(0.49)	(0.43)	(0.44)	(0.50)
	[5,170]	[30,221]				
% (SC+ST+Available Seats)	38.59	52.56	13.96***	3.22***	2.02***	-2.23***
	(20.25)	(34.99)	(0.49)	(0.39)	(0.40)	(0.47)
	[5,183]	[30,220]				
Panel B: Movements within shrug code						
% (SC+ST)	28.21	56.37	28.17***	20.02***	19.96***	17.03***
	(21.69)	(33.13)	(0.48)	(0.40)	(0.41)	(0.47)
	[5,183]	[30,221]				
% (SC+ST-Used Seats)	23.33	57.24	33.92***	26.24***	26.21***	25.98***
	(23.02)	(32.84)	(0.49)	(0.43)	(0.44)	(0.51)
	[5,171]	[30,221]				
% (SC+ST+Available Seats)	36.09	52.08	16.00***	5.41***	5.27***	-11.20***
	(20.89)	(36.03)	(0.49)	(0.42)	(0.43)	(0.47)
	[5,183]	[30,170]				

Notes: %(SC+ST) is the percentage of Grade 1 enrollment taken by students who are from a Scheduled Caste or a Schedule Tribe. %(SC+ST-Used Seats) estimates the percentage of Grade 1 enrollment taken by students who are from a Scheduled Caste or a Schedule Tribe in the absence of the RTE policy assuming all used RTE seats are taken by children from those groups and that these students would otherwise go to a public school. %(SC+ST+Available Seats) estimates the percentage of Grade 1 enrollment taken by students who are from a Scheduled Caste or a Schedule Tribe if all (additionally available) RTE seats are taken by children from those groups (and these students come from public schools). Column 1 shows the mean in private schools (standard deviation in parenthesis), while Column 2 shows the mean in public schools (standard deviation in parenthesis). Column 3 presents the difference (with its standard error in parenthesis), Column 4 presents the difference with block fixed effects (with its standard error in parenthesis), Column 5 presents the difference with postal code fixed effects (with its standard error in parenthesis), Column 6 presents the difference with village/town fixed effects, as defined by the SHRUG-ID created by Asher et al. (2021) (with its standard error in parenthesis). The estimates are weighted by total enrollment. Panel A assumes all movements of students happen within postal codes, while Panel B assumes they happen within villages/towns, as defined by the SHRUG-ID (Asher et al., 2021). The data is restricted to schools with at least some enrollment in Grade 1. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

E Tables and figures controlling for students' preferences

E.1 Main tables

Table E.1: Balance across lottery winners and losers, controlling for students' preferences

	Admin data		Phone survey #1		Phone survey #2	
	Control mean (1)	Treatment differential (2)	Control mean (3)	Treatment differential (4)	Control mean (5)	Treatment differential (6)
Female	0.49	0.00	0.50	-0.02	0.49	-0.02
	(0.50)	(0.01)	(0.50)	(0.02)	(0.50)	(0.02)
	[4,932]	[10,079]	[2,222]	[3,831]	[1,203]	[2,057]
Age (Jan 1st, 2019)	4.06	-0.01*	4.00	-0.02**	3.98	-0.02
	(0.93)	(0.01)	(0.92)	(0.01)	(0.89)	(0.02)
	[4,932]	[10,079]	[2,222]	[3,831]	[1,203]	[2,057]
Scheduled Caste	0.17	-0.00	0.17	0.01	0.16	0.01
	(0.38)	(0.01)	(0.37)	(0.01)	(0.37)	(0.02)
	[4,932]	[10,079]	[2,222]	[3,831]	[1,203]	[2,057]
Scheduled Tribe	0.17	-0.00	0.12	-0.00	0.10	0.01
	(0.38)	(0.01)	(0.32)	(0.01)	(0.30)	(0.01)
	[4,932]	[10,079]	[2,222]	[3,831]	[1,203]	[2,057]
Other Backward Class	0.54	-0.00	0.58	-0.00	0.60	-0.01
	(0.50)	(0.01)	(0.49)	(0.02)	(0.49)	(0.02)
	[4,932]	[10,079]	[2,222]	[3,831]	[1,203]	[2,057]
Rural	0.40	-0.00	0.32	0.00	0.31	0.00
	(0.49)	(0.00)	(0.46)	(0.00)	(0.46)	(0.00)
	[4,932]	[10,079]	[2,222]	[3,831]	[1,203]	[2,057]
Surveyed			0.44	0.02**	0.26	0.03***
			(0.50)	(0.01)	(0.44)	(0.01)
			[4,932]	[10,079]	[4,932]	[10,079]
Allocated a seat			0.21	0.74***	0.17	0.78***
			(0.41)	(0.01)	(0.38)	(0.02)
			[2,173]	[3,796]	[1,088]	[2,138]

Notes: Odd columns contain the control (lottery losers) mean, standard deviation of the mean (in parentheses), and the number of observations in the control group (in square brackets). Even columns report the treatment effect (difference between lottery winners and losers), the standard error of the effect (in parentheses), and the number of observations in the treatment group (in square brackets). Columns 1–2 focus on the full sample. The p-value of the null hypothesis that the differences across all the observable applicant characteristics (Column 2) are jointly zero is .81. Columns 3–4 focus on those who answered our first phone survey. The p-value of the null hypothesis that the differences across all the observable applicant characteristics (Column 4) are jointly zero is .25. Columns 5–6 focus on those who answered our second phone survey. The p-value of the null hypothesis that the differences across all observable applicant characteristics (Column 6) are jointly zero is .62. All treatment estimates control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.2: Effect on the extensive margin of enrollment, controlling for students' preferences

	Any school				Private school			
	Control mean	ITT	CCM	LATE	Control mean	ITT	CCM	LATE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All	0.86	0.14***	0.83	0.19***	0.82	0.18***	0.78	0.24***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
		[6,053]		[5,969]		[6,002]		[5,928]
Nursery	0.81	0.19***	0.77	0.25***	0.79	0.21***	0.74	0.29***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
		[3,103]		[3,062]		[3,070]		[3,035]
Kindergarten	0.87	0.13***	0.85	0.17***	0.82	0.18***	0.79	0.24***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
		[1,766]		[1,741]		[1,756]		[1,735]
Grade 1	0.98	0.02**	0.97	0.03**	0.91	0.09***	0.89	0.12***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
		[1,184]		[1,166]		[1,176]		[1,158]

Notes: Columns 1 and 5 report the control (lottery losers) mean and the standard error of the mean (in parentheses). Columns 2 and 6 list the itent-to-treat (ITT) effect (difference between lottery winners and losers), the standard error of the effect (in parentheses), and the number of observations used to estimate the effect (in square brackets). Columns 3 and 7 report the control complier mean (CCM) — the mean outcomes for lottery loser compliers — and the standard error of the CCM (in parentheses). Columns 4 and 8 list the local average treatment effect (LATE) of being assigned an RTE seat (instrumented by winning the lottery), the standard error of the effect (in parentheses), and the number of observations used to estimate the effect (in square brackets). All treatment estimates control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.3: Effect on the characteristics of the school a child attends, controlling for students' preferences

	English medium (1)	% students ST & SC (2)	Facility index (3)	Enrollment (4)	Teachers (5)	PTR (6)
Panel A: ITT						
Lottery seat	.054*	.17	-.004	50**	1.1**	2.2**
	(.03)	(.63)	(.037)	(20)	(.58)	(1.1)
N. of obs.	1,166	818	820	777	799	750
Control mean	0.56	28.69	0.71	403.56	13.35	29.93
Control mean | enrolled	0.57	29.34	0.73	413.53	13.66	30.67
% Enrolled (Control)	98.71	97.78	97.78	97.59	97.73	97.59
Panel B: LATE						
Allocated an RTE seat	.076*	.44	-.017	61**	1.4*	3.5**
	(.042)	(.87)	(.051)	(27)	(.78)	(1.6)
N. of obs.	1,151	810	812	769	791	742
CCM	0.51	29.33	0.72	443.22	14.57	28.75
CCM | enrolled	0.52	29.94	0.74	455.59	14.97	29.45
% Enrolled (CCM)	97.45	97.76	97.79	97.63	97.62	97.53

Notes: Panel A presents the ITT effects of winning a seat through the lottery on different characteristic of the school the child is enrolled in. Panel B presents the LATE of being allocated an RTE (instrumenting with the outcome of the lottery) on different characteristics of the school the child is enrolled in. CCM denotes the mean outcomes for lottery loser compliers. In Column 1, the outcome is whether the child attends an English medium schools or not. In Column 2, the outcome is the percentage of enrollment taken by Scheduled Castes and Tribes in the school the child attends. In Column 3, the outcome is a principal component analysis (PCA) facility index based on whether the school has computer assisted learning, a homeroom, electricity, a library, a playground, a solid building, a boundary wall, functioning toilets, and solid classrooms. In Columns 4-6 the outcomes are enrollment, number of teachers, and the pupil-teacher ratio (PTR). All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.4: Effect on fees, controlling for students' preferences

	INR			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Panel A: ITT				
Lottery seat	3,281***	4,900***	2,151***	1,470***
	(215)	(351)	(328)	(379)
Control mean	5,628	5,826	4,869	6,456
Control mean in private	7,615	9,240	6,249	7,294
% out of school (control)	20	33	15	2
% in public (control)	6.3	3.9	7.2	9.5
N. of obs.	4,499	2,171	1,454	874
Panel B: LATE				
Allocated an RTE seat	4,278***	6,185***	2,852***	2,010***
	(290)	(482)	(446)	(501)
CCM	5,470	5,983	4,468	5,955
CCM in private	7,927	10,059	6,011	7,019
% out of school (CCM)	17	23	15	2.5
% in public (CCM)	4.7	2.7	6	8.4
N. of obs.	4,469	2,161	1,443	865

Notes: Fee information comes from administrative data. Students in public schools or not enrolled in school are assigned zero fees. Panel A presents the ITT effect of winning a lottery seat. Panel B presents the LATE of being allocated an RTE (instrumenting with the outcome of the lottery) on the market price of the school a child attends. All regressions control for "full preference" list fixed effects. CCM denotes the mean outcomes for lottery loser compliers. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.5: Effect on enrollment in top choice controlling for students' preferences

	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Lottery seat at first choice	.53***	.6***	.47***	.41***
	(.014)	(.018)	(.024)	(.034)
N. of obs.	5,360	2,782	1,605	973
Control mean	0.42	0.35	0.47	0.53
Control mean | enrolled	0.49	0.44	0.55	0.53
Control mean | enrolled & no RTE seat	0.57	0.57	0.57	0.59
% Enrolled (Control)	84.67	79.01	86.09	98.23
% RTE seat (Control)	26.78	29.32	23.25	25.23

Notes: This table presents the ITT effects of winning a place in the first-choice school through the lottery on the likelihood of enrolling in this top-choice school. All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.6: Effect on government expenditure, controlling for students' preferences

	INR			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Market price				
Allocated an RTE seat	4,278***	6,185***	2,852***	2,010***
	(290)	(482)	(446)	(501)
CCM	5,470	5,983	4,468	5,955
CCM in private	7,927	10,059	6,011	7,019
% out of school (CCM)	17	23	15	2.5
% in public (CCM)	4.7	2.7	6	8.4
N. of obs.	4,469	2,161	1,443	865
Panel B: Reimbursed fee				
Allocated an RTE seat	6,008***	6,761***	5,149***	5,636***
	(71)	(99)	(112)	(153)
N. of obs.	4,469	2,161	1,443	865
Panel C: Non-limit reimbursed fee				
Allocated an RTE seat	9,748***	12,169***	7,319***	7,965***
	(308)	(504)	(469)	(527)
N. of obs.	4,469	2,161	1,443	865

Notes: Fee information comes from administrative data. Students in public schools or not enrolled in school are assigned zero fees. Panel A presents the LATE of being allocated an RTE seat (instrumenting with the outcome of the lottery) on the market price of the school a child attends. Panel B presents the LATE of being allocated an RTE seat (instrumenting with the outcome of the lottery) on the reimbursed fee (set to zero for children without an RTE seat). Panel C presents the LATE of being allocated an RTE (instrumenting with the outcome of the lottery) on the hypothetical reimbursed fee in the absence of the maximum reimbursement limit (set to zero for children without an RTE seat). All regressions control for "full preference" list fixed effects. CCM denotes the mean outcomes for lottery loser compliers. Table E.14 presents the ITT estimates of winning a lottery seat. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

E.2 Appendix tables

Table E.7: Attrition by child characteristics, controlling for students' preferences

	Survey #1	Survey #2
	(1)	(2)
Female	.0042	.005
	(.0083)	(.0073)
Age (Jan 1st, 2019)	-.015	-.0097
	(.013)	(.011)
Scheduled Caste	-.037**	-.024
	(.018)	(.016)
Scheduled Tribe	-.074***	-.065***
	(.019)	(.017)
Other Backward Class	-.012	.0011
	(.016)	(.014)
Rural	.042	.031
	(.038)	(.033)
N. of obs.	15,011	15,011
Outcome mean	.43	.25

Notes: The outcome is whether we were able to conduct the interview (=1). All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.8: Compliance, controlling for students' preferences

	Allotted an RTE seat			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Allocated a seat	.77***	.78***	.76***	.77***
	(.011)	(.015)	(.021)	(.026)
N. of obs.	5,969	3,062	1,741	1,166
Control mean	0.17	0.18	0.17	0.18

Notes: This table presents the effect of winning a lottery seat on being allotted an RTE seat. All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.9: Effect on the extensive margin of enrollment, controlling for the probability of being assigned to a private school: Lee bounds and stratas with low attrition controlling for students' preferences

	Strata without attrition		Low attrition strata			Lee bounds	
	ITT	LATE	Differential attrition	ITT	LATE	ITT	
						LB	UB
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: All grades							
Private school (19-20)	0.16***	0.20***	0.00	0.19***	0.24***	0.12	0.26
	(0.04)	(0.05)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
	[367]	[362]	[6,294]	[3,104]	[3,070]	[2,913]	[2,913]
Any school (19-20)	0.11***	0.14***	0.00	0.14***	0.18***	0.08	0.19
	(0.03)	(0.04)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
	[240]	[236]	[6,294]	[3,138]	[3,097]	[2,937]	[2,937]
Panel B: Nursery							
Private school (19-20)	0.18***	0.22***	-0.00	0.22***	0.28***	0.12	0.27
	(0.05)	(0.06)	(-0.00)	(0.02)	(0.02)	(0.02)	(0.03)
	[146]	[144]	[2,983]	[1,543]	[1,529]	[1,420]	[1,420]
Any school (19-20)	0.16***	0.20***	-0.00	0.19***	0.25***	0.11	0.24
	(0.04)	(0.05)	(-0.00)	(0.02)	(0.02)	(0.02)	(0.03)
	[150]	[146]	[2,983]	[1,567]	[1,549]	[1,436]	[1,436]
Panel C: Kindergarten							
Private school (19-20)	0.11	0.12	-0.01	0.16***	0.21***	0.13	0.34
	(0.10)	(0.11)	(-0.01)	(0.02)	(0.02)	(0.03)	(0.04)
	[20]	[20]	[2,098]	[974]	[961]	[914]	[914]
Any school (19-20)	.	.	-0.01	0.12***	0.15***	0.09	0.23
	(.)	(.)	(-0.01)	(0.02)	(0.02)	(0.02)	(0.03)
	[.]	[.]	[2,098]	[981]	[965]	[920]	[920]
Panel D: Grade 1							
Private school (19-20)	0.13*	0.18*	0.02	0.11***	0.15***	0.09	0.14
	(0.06)	(0.09)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
	[68]	[68]	[1,213]	[584]	[577]	[579]	[579]
Any school (19-20)	0.05	0.07	0.02	0.03**	0.04**	0.01	0.02
	(0.04)	(0.06)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
	[70]	[70]	[1,212]	[588]	[581]	[581]	[581]

Notes: Columns 1–2 report the results restricting the sample to strata without attrition. Column 1 shows the ITT effect of winning the lottery, and Column 2 the LATE of being assigned an RTE seat (instrumented with winning the lottery). Columns 3–5 show the results after dropping the 25% of the strata with the most differential attrition. Column 3 shows the results of the differential attrition, Column 4 the ITT effect, and Column 5 the LATE of being assigned an RTE seat. Columns 6–7 show Lee (2009) style bounds — Column 6 has the lower bound (LB), while Column 7 has the upper bound for (UB) — for the ITT effect of winning the lottery. Standard errors are in parentheses. The number of observations in the treatment effects estimates is in square brackets. All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.10: Heterogeneity on school enrollment ITT, controlling for the probability of being assigned to a private school and students' preferences

	Any school (19-20)		Private school (19-20)	
	All	Grd 1	All	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Heterogeneity by gender				
Lottery seat	.13***	.014	.18***	.098***
	(.011)	(.011)	(.013)	(.023)
Female	-.0046	-.007	-.00031	.02
	(.015)	(.012)	(.017)	(.027)
Lottery seat × Female	.0067	.008	.0018	-.025
	(.015)	(.012)	(.017)	(.028)
N. of obs.	6,053	1,184	6,002	1,176
Control mean	.87	.99	.83	.92
Panel B: Heterogeneity by parental education				
Lottery seat	.14***	.02**	.19***	.089***
	(.0091)	(.009)	(.01)	(.017)
Mother HS	.045**	.012	.06**	.045
	(.021)	(.0076)	(.023)	(.037)
Lottery seat × Mother HS	-.05**	-.019**	-.068***	-.043
	(.022)	(.0096)	(.025)	(.042)
N. of obs.	5,858	1,152	5,812	1,143
Control mean	.87	.99	.83	.92
Panel C: Heterogeneity by caste				
Lottery seat	.14***	.015	.17***	.074**
	(.019)	(.015)	(.022)	(.034)
Other Backward Class (OBC)	.0025	.016	-.0041	.033
	(.02)	(.011)	(.023)	(.034)
Scheduled Tribe (ST)	-.016	.0043	-.014	.02
	(.028)	(.014)	(.032)	(.043)
Scheduled Caste (SC)	-.041	-.045	-.073**	-.11**
	(.026)	(.03)	(.03)	(.053)
Lottery seat × OBC	-.016	-.013	-.0099	-.024
	(.021)	(.013)	(.024)	(.036)
Lottery seat × ST	-.0023	-.001	.0024	-.0081
	(.028)	(.016)	(.033)	(.048)
Lottery seat × SC	.02	.046	.053*	.12**
	(.027)	(.033)	(.031)	(.057)
N. of obs.	6,053	1,184	6,002	1,176
Control mean	.87	.99	.83	.92

Notes: This table presents the ITT estimates of being assigned a seat by winning the lottery. The outcome in Columns 1–2 is whether the child was enrolled in any school in 2019–2020 (=1). The outcome in Columns 3–4 is whether the child was enrolled in a private school in 2019–2020 (=1). Mother HS indicates whether the mother completed high school. Columns 1 and 3 use the full sample, while Columns 2 and 4 use only Grade 1 students. All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.11: Heterogeneity on school enrollment LATE, controlling for the probability of being assigned to a private school and students' preferences

	Any school (19-20)		Private school (19-20)	
	All	Grd 1	All	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Heterogeneity by gender				
Allocated an RTE seat	.18***	.02	.24***	.14***
	(.015)	(.015)	(.017)	(.033)
Female	-.013	-.009	-.0069	.028
	(.019)	(.016)	(.022)	(.037)
Allocated an RTE seat × Female	.013	.0095	.0061	-.037
	(.02)	(.017)	(.023)	(.04)
N. of obs.	5,969	1,166	5,928	1,158
CCM	.83	.97	.78	.89
Panel B: Heterogeneity by parental education				
Allocated an RTE seat	.2***	.028**	.26***	.12***
	(.013)	(.013)	(.014)	(.024)
Mother HS	.063**	.02*	.085***	.067
	(.027)	(.011)	(.03)	(.048)
Allocated an RTE seat × Mother HS	-.07**	-.027**	-.096***	-.068
	(.029)	(.014)	(.033)	(.057)
N. of obs.	5,783	1,136	5,743	1,127
CCM	.83	.97	.78	.89
Panel C: Heterogeneity by caste				
Allocated an RTE seat	.19***	.025	.23***	.12**
	(.026)	(.024)	(.029)	(.055)
Other Backward Class (OBC)	.001	.02	-.013	.047
	(.026)	(.017)	(.029)	(.049)
Scheduled Caste (SC)	-.051	-.054	-.095**	-.13*
	(.034)	(.04)	(.039)	(.07)
Scheduled Tribe (ST)	-.032	.0045	-.032	.018
	(.037)	(.021)	(.042)	(.062)
Allocated an RTE seat × OBC	-.016	-.021	-.0015	-.056
	(.028)	(.021)	(.032)	(.057)
Allocated an RTE seat × SC	.035	.057	.08*	.14*
	(.037)	(.046)	(.042)	(.081)
Allocated an RTE seat × ST	.016	-.004	.023	-.02
	(.04)	(.025)	(.046)	(.072)
N. of obs.	5,969	1,166	5,928	1,158
CCM	.83	.97	.78	.89

Notes: This table presents the LATE of being assigned an RTE place (instrumented by winning the lottery). CCM denotes the mean outcomes for lottery loser compliers. The outcome in Columns 1–2 is whether the child was enrolled in any school in 2019–2020 (=1). The outcome in Columns 3–4 is whether the child was enrolled in a private school in 2019–2020 (=1). Mother HS indicates whether the mother completed high school. Columns 1 and 3 use the full sample, while Columns 2 and 4 use only Grade 1 students. All regressions control for "full preference" list fixed effects. Table E.10 provides the ITT effect of winning a lottery seat. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.12: Effect of winning different lottery seats on enrollment in the top-choice school, controlling for students' preferences

	(1)	(2)	(3)	(4)	(5)	(6)
Won lottery	.48***					
	(.014)					
Won seat in first choice		.53***	.72***	.66***	.7***	.66***
		(.014)	(.023)	(.026)	(.039)	(.041)
Won seat in second choice				-.28***	-.26***	-.29***
				(.036)	(.069)	(.07)
Won seat in third choice						-.29***
						(.057)
N. of obs.	5,360	5,360	1,461	1,461	555	555
Sample	Full	Full	≥ 2 choices	≥ 2 choices	≥ 3 choices	≥ 3 choices

Notes: This table presents the effect of winning different lottery seats on the likelihood of enrolling in the top-choice school. All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.13: Effect on the diversity of the student body, controlling for students' preferences

	% SC (1)	% ST (2)	% SC+ST (3)
Panel A: ITT			
Lottery seat	.055	-1	-.99
	(.81)	(.96)	(1.3)
Scheduled Tribe	-.11	-1.1	-1.3
	(.9)	(1.4)	(1.8)
Scheduled Caste	2.8**	-.15	2.7
	(1.3)	(1)	(1.9)
Other Backward Class	-.62	-.91	-1.5
	(1)	(1)	(1.6)
Lottery seat × Scheduled Tribe	.13	1.3	1.5
	(.9)	(1.4)	(1.7)
Lottery seat × Scheduled Caste	-1.8	.14	-1.6
	(1.3)	(1.1)	(1.9)
Lottery seat × Other Backward Class	1.1	1.2	2.2
	(1)	(1)	(1.5)
N. of obs.	818	818	818
Control mean	12.82	15.87	28.69
Control mean | enrolled	13.11	16.23	29.34
% Enrolled (Control)	97.78	97.78	97.78
Panel B: LATE			
Allocated an RTE seat	.37	-1.5	-1.2
	(1.3)	(1.6)	(2.1)
Allocated an RTE seat × Scheduled Caste	-2.7	.31	-2.4
	(1.9)	(1.7)	(2.8)
Allocated an RTE seat × Scheduled Tribe	-.17	1.9	1.8
	(1.4)	(2.2)	(2.7)
Allocated an RTE seat × Other Backward Class	1.3	1.7	3
	(1.5)	(1.6)	(2.3)
Scheduled Caste	3.6*	-.27	3.3
	(1.9)	(1.6)	(2.7)
Scheduled Tribe	.12	-1.7	-1.5
	(1.3)	(2.1)	(2.7)
Other Backward Class	-.9	-1.4	-2.3
	(1.4)	(1.5)	(2.2)
N. of obs.	810	810	810
CCM	12.72	16.62	29.33
CCM | enrolled	12.99	16.94	29.94
% Enrolled (CCM)	97.76	97.76	97.76

Notes: Panel A presents the ITT effects of winning a seat through the lottery on the proportion of students from Scheduled Castes (SC) and Scheduled Tribes (ST). Panel B presents the LATE of being allocated an RTE (instrumenting with the outcome of the lottery) on the proportion of students from SC and ST. CCM denotes the mean outcomes for lottery loser compliers. All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table E.14: Intent-to-treat effect on government expenditure, controlling for students' preferences

	INR			
	All	NU	KG	Grd 1
	(1)	(2)	(3)	(4)
Panel A: Market price				
Lottery seat	3,281***	4,900***	2,151***	1,470***
	(215)	(351)	(328)	(379)
Control mean	5,628	5,826	4,869	6,456
Control mean in private	7,615	9,240	6,249	7,294
% out of school (control)	20	33	15	2
% in public (control)	6.3	3.9	7.2	9.5
N. of obs.	4,499	2,171	1,454	874
Panel B: Reimbursed fee				
Lottery seat	4,580***	5,349***	3,789***	4,174***
	(84)	(121)	(132)	(193)
N. of obs.	4,499	2,171	1,454	874
Panel C: Non-limit reimbursed fee				
Lottery seat	7,431***	9,615***	5,404***	5,884***
	(243)	(384)	(368)	(465)
N. of obs.	4,499	2,171	1,454	874

Notes: Fee information comes from administrative data. Students in public schools or not enrolled in school are assigned zero fees. Panel A presents the ITT effects of being allocated an RTE through the lottery on the market price of the school a child attends. Panel B presents the ITT effects of being allocated an RTE through the lottery on the reimbursed fee (set to zero for children without an RTE seat). Panel C presents the ITT effects of being allocated an RTE through the lottery on the hypothetical reimbursed fee in the absence of the maximum reimbursement limit (set to zero for children without an RTE seat). All regressions control for "full preference" list fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.