Essays in Development Economics

By

Monica Agarwal

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Agricultural and Applied Economics)

at the
UNIVERSITY OF WISCONSIN–MADISON
2024

Date of final oral examination: 06/20/2024

The dissertation is approved by the following members of the Final Oral Committee:

Laura Schechter, Professor, Agricultural and Applied Economics
Priya Mukherjee, Assistant Professor, Agricultural and Applied Economics
Jeffrey Smith, Professor, Economics
Andrew Stevens, Assistant Professor, Agricultural and Applied Economics
Rhiannon Jerch, Assistant Professor, Agricultural and Applied Economics

Dedication

Dedicated to my father Alok Agarwal, and my mother Sapna Agarwal.

Acknowledgments

I am forever grateful to my parents, Alok Agarwal and Sapna Agarwal for being my biggest pillars of strength, for their unconditional love and for supporting me in all my endeavors. My grandparents, Kripa Shanker and Sheela Agarwal for their abundant love and for always showering me with blessings. My spouse and my best friend, Moshi Alam for standing by my side through thick and thin and for being my sounding board. My brother, Amol Agarwal and sister-in-law, Manisha Saraf for always cheering me on. My most adorable niece, Ivana Agarwal who brightens my day without fail. My parents-in-law, Sadrul Alam and Mita Alam for their love and support. I am truly grateful to have such a wonderfully supportive family and I thank them for supporting me in this journey.

My time here in the Department of Agricultural and Applied Economics at UW-Madison has been a very enriching experience. I am truly grateful to my incredible advisors Laura Schechter, Jeffrey Smith and Priya Mukherjee for their invaluable guidance. I truly admire them for their mentorship and for their constant efforts in advancing the careers of their students. I genuinely believe that I won the advisor lottery and could not have asked for a better PhD committee. Their unwavering support and insightful feedback have been instrumental in shaping my research and academic journey. My heartfelt thanks to Professor Bradford Barham, with whom I worked in the early years of my PhD, for his encouragement and insights. My gratitude also goes to Leah Bevis, Ian Coxhead, Steven Deller, Paul Dower, Jeremy Foltz, Tarun Jain, Rhiannon Jerch, Sarah

Johnston, Steven Lehrer, Ian McCarthy, Dominic Parker, Nishith Prakash, Andrew Stevens, and Emilia Tjernström who helped me at various stages. I am also very thankful to my peers, friends and colleagues whom I met during my PhD and learned a lot from - Sossou Adjisse, Priyadarshi Amar, Gary Baker, Eduardo Cenci, Pukitta Chunsuttiwat, Gaurav Doshi, Felipe Parra Escobar, Vikas Gawai, Manisha Jain, Nicole Karwowski, Erik Katovich, Amrita Kulka, Itzel De Haro Lopez, Dennis McWeeny, Ana Paula Melo, Osas Olurotimi, Pedro Magana Saenz, Elan Segarra, Kait Sims, Talita Silva, Mizuhiro Suzuki, and Karin Wu. You all made this journey memorable and rewarding. Many thanks also to the department staff, especially Eric Dieckman, Erin Wall, and Mary Trevelen for helping me navigate through all the administrative aspects of the program during all these years.

I am deeply appreciative of the grant support I received from J-PAL and a Dissertation Research Grant from AAE which made it possible for me to collect data for my research. My thanks to the survey firm NEERMAN, Nandish Kenia and the team of enumerators who conducted excellent fieldwork as part of the data collection process.

Finally, I would like to thank all my wonderful friends for all the fun memories. My thanks to Arkadipta Bakshi, Akarshik Banerjee, Shubhashrita Basu, Ashika Bhargav, Saloni Bhogale, Trisha Chanda, Sohang Kundu, Rini Tarafdar, Mayukh Talukdar, and Shreya Singh. Your friendship has been a constant source of comfort and happiness throughout this journey.

To everyone who has been a part of this incredible journey, your support and encouragement has been invaluable. Thank you from the bottom of my heart.

Contents

Introduction

In this dissertation, I contribute to the understanding of topics in the field of the economics of education and development economics in the context of India. The first two chapters focus on the economics of education and study India's affirmative action policy, the Right to Education Act, and throw light on its effectiveness on the lives of disadvantaged children. The third chapter focuses on development economics and studies how women's inheritance rights might improve their socioeconomic outcomes.

In the first chapter, I study the causal impact of India's Right to Education Act (RTE) on educational outcomes of children. As one the largest affirmative action policies in the world, it targets millions of children of school entry age and mandates all private schools to provide free of cost schooling to low socioeconomic status children by reserving seats for them in entry level grades. I find that as a result of winning entry to any private school under this policy, children's educational outcomes improve and test scores increase by a significant margin. Furthermore, attending higher quality private schools improve test scores of beneficiaries even further, relative to lower quality private schools. My findings show that there is considerable heterogeneity within the private schooling sector and that a single estimate of the private school premium masks the underlying heterogeneity within the sector - an important contribution in the literature. This study is from the context of remote learning during the COVID-19 induced school closures, and provides evidence that private schools are effective not just during in-person settings but also in remote learning contexts - another important contribution, since most of our understanding about private school effectiveness is from in-person learning contexts.

In the second chapter, I examine if a positive schooling shock for one child in the household leads to any spillover effects on the educational outcomes of their sibling in the context of RTE lotteries. My findings indicate that spillover effects exist and vary by the age of siblings. Younger siblings in households that win the grade 1 RTE lottery for the applicant child are less likely to be formally enrolled in school compared to their peers in losing households, but benefit from increased access to remote learning resources provided by the older applicant child's private school. Additionally, there are no significant differences in parental monetary and time investments between siblings in winning and losing households. This study shows that well-implemented affirmative action policies during economic hardships, such as the COVID-19 pandemic, can act as a safety net not only for the targeted individuals, but also benefit non-targeted individuals and can mitigate long-term educational inequalities.

In the third chapter of my dissertation, I along with my coauthor study whether improving women's inheritance rights that increase their access to household property, affect their outcomes in their marital household. We estimate the causal impact of amendments to the Hindu Succession Act, which aimed to improve women's property inheritance rights in India, on the likelihood of the presence of a toilet in their marital households using a difference-in-differences framework, and allowing for dynamic and heterogeneous treatment effects. Our findings show that the policy had a positive impact on the presence of a toilet in treated women's marital households, with the effects being concentrated in the states that adopted the policy late. We attribute our results to the policy's role in increasing women's years of educational attainment and weakly increasing their intra-household bargaining power.

In conclusion, this dissertation contributes to the understanding of key issues in the economics of education and development economics in the Indian context. Together, these studies provide valuable insights for policymakers aiming to design effective educational and social policies that promote equity and development.

Chapter 1

The Role of Affirmative Action in Enrollment, Test Scores, and School Quality: Evidence from India

1.1 Abstract

Worldwide, affirmative action policies are implemented as a means to promote social equity. India's Right to Education Act (RTE), one of the largest affirmative action policies in the world, mandates all private schools to reserve 25% of incoming seats at entry-level grades for low socioeconomic status students. Despite being in existence for more than a decade, the effectiveness of this policy remains understudied. In this paper, I estimate the causal impact of RTE's 25% quotas on children's learning outcomes using a combination of rich administrative and survey data in a large state in India. I leverage the lottery-based allocation of oversubscribed schools to identify the causal impact of being a beneficiary under this policy. I find that the policy improves children's English test scores by 0.18 SD via beneficiaries attending better schools, and investing more time in educational activities. Furthermore, while the policy allocates children to private schools, there exists a large variation in school quality within the private sector. Motivated by the existence of this within-sector heterogeneity in quality, I uncover the distribution of effects within the private sector, and find

that higher quality private schools boost English test scores by 0.5-0.7 SD, relative to their lower quality counterparts. My findings are from a context when all learning is remote, and indicate that private schools, especially the ones at the upper end of the quality distribution, do a better job at adapting to, and implementing remote educational technologies, and in doing so, they also enhance children's learning.

1.2 Introduction

Governments across the world implement affirmative action policies as a means to promote social equity. Such policies aim to redress long histories of discrimination against historically disadvantaged groups. The majority of such policies typically focus on later life stages of individuals, such as college admissions, or the workplace. However, how might individuals' life trajectories change if such disparities are reduced early in life? A growing literature suggests that reducing disparities early in life is important for the formation of cognitive skills and that later life interventions may be too late to achieve this in a cost-effective way, for example, Cunha, Heckman and Schennach (2010). As one of the world's largest affirmative action policy that targets children of school entry age, India's Right to Education Act (RTE) provides a unique opportunity to study this question.

The RTE mandates all private schools in India to reserve seats for disadvantaged children at entry level grades, with the goal of reducing segregation within classrooms. The scale of the policy is huge - in 2018-19 alone, the policy benefited approximately 4 million children, and has the potential to impact about 16 million children, if implemented nationally (Indus Action, 2019; Romero and Singh, 2023). As a direct effect, the policy improves accessibility to private schools for economically disadvantaged classes. Thus, a first order question is to study the effectiveness of this policy, or in other words, the impact of attending a private school under the policy.

In addition, the private schooling sector in India has been steadily growing and accounted for 45% of the primary grade enrollment in 2020.[1] Given the rapid growth in the market share of fee-charging private schools, both at the upper and lower end of the quality distribution, school effectiveness is likely to vary within the private sector. This in turn, signals the importance of examining the distribution of effects *within* the private sector.

Hence, in this paper I study two main questions. First, I ask: does being a beneficiary under the RTE quotas improve disadvantaged children's educational outcomes? Second, do the effects of this policy vary by the quality of private schools that beneficiaries attend?

I study the impact of this policy in the context of Maharashtra, the second most populous state in India. Under the policy, private schools are mandated to reserve up to 25% of the incoming seats at entry-level grades for disadvantaged children. Allocation of private school seats to applicants under the policy is based on a lottery mechanism which ensures that applicants who submit the same school preferences, and live in the same neighborhood, have an equal chance of winning a seat at any given school that they listed in their application. Those who win entry to private schools under this policy are eligible to get tuition-free education from these schools until they finish grade 8, with the government reimbursing the schools up to a cap. The outside option for those who lose under the policy, is to either attend a private school of their choice as a fee-paying student, attend a government school (which are free of cost, but often of lower quality), or remain out-of-school. I use the feature of lottery-induced allocation of oversubscribed private school seats to estimate the causal impact of the RTE policy on children's educational outcomes within an instrumental variables framework. Methodologically, I follow the recent methods by Abdulkadiroğlu et al. (2017) and utilize the within-variation in lottery outcomes of applicants who had a similar simulated ex-ante probability of winning the private school lottery under the allocation mechanism, which takes into account the school preferences submitted at the time of application.

To do this I use the administrative data of the population of children who applied for grade 1 private school lotteries under RTE's 25% quotas, in the 2020-21 school year. I supplement

[1] World Bank data (2020)

this with data from a phone survey which I designed and administered with a sample of applicant households, to collect detailed information on children's education, schooling, and performance on phone-based assessments in English and Math.[2] This gives me a sample of 2329 applicant households for whom I have a rich data on household characteristics, children's schooling, their performance on phone-based assessments, their time-use, parental investments, and school inputs.

The data, however, corresponds to the period of COVID-19 induced school closures. Like many low- and middle-income countries, pandemic-induced school closures lasted for a long time in India. While schools were closed for in-person instruction, the majority of schools transitioned to various forms of remote instruction (both asynchronous and synchronous) at some point during the 2020-21 school year. Since the majority of current evidence on private school effectiveness is from in-person learning contexts, this setting provides me a unique opportunity to study whether private schools are effective when learning is remote. Thus, the findings in my paper are most relevant to the context of remote learning; however, it is worth noting that my results align closely with past evidence in comparable interventions, where the mode of instruction is in-person. I discuss this in more depth in the subsequent paragraphs.

My findings show that the RTE policy led to significant improvements in educational outcomes of children. One and a half years after exposure to RTE, I find that quota children who won the private school lottery were much more likely to be enrolled in school in the two academic years 2020-21 and 2021-22.[3] The effect sizes are substantial - for compliers, the likelihood of being enrolled at any school increases by 13.3 and 4.6 percentage points, in 2020-21 and 2021-22, respectively. Given that primary school enrollment is near universal in India, these increases in enrollment largely reflect that the RTE was useful in insuring disadvantaged

[2]This assessment was adapted from the phone-based learning instruments used by Romero and Singh (2023) and Angrist, Bergman and Matsheng (2020).

[3]Applications for private school admissions under the RTE 25% quotas were made for the 2020-21 academic year, and I conducted phone-surveys with a sample of these applicants during the middle of the following academic year i.e., 2021-22. This allows me to study their enrollment decision in these two academic years.

children against the risk of non-enrollment during a period of massive disruptions to learning. The gains however, are not just limited to enrollment, but also include gains in test scores of children - for the compliers, being a quota student at a private school improves performance in English by 0.18 SD (p-value < 0.05). There is evidence of suggestive gains in Math (by 0.14 SD), however, the impact on Math is not statistically distinguishable from zero. Even though my findings come from a remote learning context, they are strikingly similar to prior estimates of private school effectiveness when learning happens in-person. Muralidharan and Sundararaman (2015) find gains of 0.12 SD units in English, but none in Math, for winners of private school vouchers in India after 4 years.[4] This suggests that private schools are effective not just during in-person settings, but also when learning is remote.

In order to interpret these results, it is helpful to learn about the composition of the counterfactual group. While treated compliers are a homogeneous group who attend private schools under the RTE quotas, the same is not true for the control compliers, since they have multiple outside options to choose from, such as, attending private schools as a fee-paying student, attending government schools, or being out of school. Looking at the extensive margin of the type of school being attended, I find that for compliers, the quota receipt increases the likelihood of attending a private school by 20 percentage points. However, this is over a base of 79% private school enrollment in the control group comprising non-quota students, which indicates that the outside option for those who lose is not necessarily to attend government schools. Following Abdulkadiroğlu, Angrist and Pathak (2014), an analysis of *counterfactual destinies* for control compliers highlights that about 65% of lottery losers end up at private schools as fee-paying students, and only about 20% end up at government schools. The fact that attending private schools as fee-paying students happens to be the fallback option for the majority of lottery losers, highlights the aspect of regressive

[4] Singh (2015) finds effect sizes of similar magnitudes using value-added estimates in Andhra Pradesh, India. Romero, Sandefur and Sandholtz (2020) study the impact of allocating private management bodies to existing government schools in Liberia, and find gains of 0.13 SD in language. Using data from a different state in India (Chhattisgarh), and in an earlier version of their paper, Romero and Singh (2023) look at the impact of being RTE quota student on test scores and find gains of 0.19 SD in foundational numeracy and literacy skills.

selection within eligible groups (Romero and Singh, 2023).[5]

Next, I explore three broad mechanisms to understand the channels through which gains in children's outcomes are realized - school inputs, parental inputs, and children's own time use. I find that school inputs, and children's own time use are the main channels that explain these gains. There is some evidence of parental monetary and time investments increasing as a result of winning the lottery, however, the effect sizes are small, suggesting that parental inputs explain only a small part of the story. Examining the mechanisms in detail, I find that conditional on being enrolled in school, quota students are more likely to receive remote instruction from their school in both the academic years - by 7 and 3 percentage points in 2020-21 and 2021-22, respectively. The magnitudes of these effects reflect that the schools attended by treated compliers were more efficient in adapting to remote learning during the period of school closures. In addition, they were more likely to receive synchronous online modes of instruction (by 13.6 percentage points), relative to the non-quota students who were more likely to receive asynchronous modes like text-based communication via WhatsApp, and pre-recorded audio and video clips. While these outcomes are more reflective of school characteristics that might specifically matter during the periods of remote instruction, I also examine the impact of winning the RTE private school lottery on the overall quality of the school being attended, which is reflective of quality in business-as-usual settings. I create a school quality index using Principal Component Analysis (PCA) that combines information on school infrastructure, digital facilities, and teacher qualifications, and find that relative to the non-quota students, quota students attend schools that are 0.6 SD units better in their overall school quality index. Quota students are also more likely to be enrolled in schools that have English as the primary language of instruction, teach more subjects, and have a longer school week (by 3 hours/week).

[5]Using RTE applications data from the state of Chhattisgarh in India, Romero and Singh (2023) find evidence of regressive selection under RTE by relatively better-off households among eligible groups. They find that 50% of the applicants who lose the RTE lottery for their top choice private school, end up attending the same school as a fee-paying student. They show that only 7.4% of the program spending under RTE quotas accrues to the bottom socioeconomic quintile, compared to 24.3% in the top quintile.

Next, I uncover heterogeneity within the private sector to examine if there are gains from attending higher quality schools within the private sector. I start with a simple case of defining schools as *elite* or *budget* based on two alternate measures of school quality. Focusing on the group of beneficiaries who won the RTE private school lottery, I leverage the randomization in lottery offers at elite private schools to compare the outcomes of ex-ante similar children who had a similar ex-ante probability of winning at elite private schools, but face a randomization in winning the lottery at elite versus budget private schools. Like before, I implement this using an IV-2SLS framework which uses lottery offers at elite schools as an instrument for enrollment at elite schools as a quota student, and utilizes the within-variation in lottery outcomes of children who have a similar simulated ex-ante propensity of winning the elite private school lottery, given their school preferences (Abdulkadiroğlu et al., 2017). The first measure of school quality or school eliteness is created using administrative data on each school's annual fee (that it charges to fee-paying students), and the second measure builds on the PCA based school quality index that I create using a variety of school characteristics. The two measures of school quality show substantial positive correlation indicating that schools that are elite based on the fee-measure are also likely to be elite based on the PCA-index measure. I find that attending an elite private school significantly improves English test scores by 0.48 SD (when eliteness is defined using school fee), and by 0.69 SD (when eliteness is defined using school quality index). However, there are no statistically significant impacts on Math.

As before, I examine potential mechanisms and find that while elite and budget private schools were equally likely to provide remote instruction, elite schools were more likely to provide synchronous online instruction (by 0.10 - 0.18 percentage points), and provide longer hours of class instruction (by 2.1 - 3.1 hours/week). Relative to budget private schools, elite schools are no more likely to teach conventional subjects (Math, English, etc.) but they are significantly more likely to teach additional subjects like general knowledge, arts/crafts, music and dance. In order to further understand why remote learning might be more effective for elite schools, I compare baseline characteristics of elite and budget schools, and find that elite schools are equipped with better digital technologies (for instance, access to internet; higher per-pupil quantities of laptops, desktops, and digital boards), employ teachers with higher qualifications, and have more teachers trained in computers. Another stark difference is in the caste composition of students attending elite and budget private schools. Elite schools on average have a notably less diverse student composition, and are likely to have lower proportions of children from disadvantaged caste categories. These differences provide additional evidence of heterogeneity within the private schooling sector, which might further explain differences in school effectiveness across elite private and budget schools, especially during periods of remote learning.

Taken together, my findings indicate that private schools attended by quota students were more effective in the delivery of remote schooling inputs, and enhanced children's learning during the period of school closures. They increased students' accountability by holding regular synchronous classes, providing student-teacher interaction, and keeping them engaged with school activities for more hours per week. In addition, I find evidence of substantial heterogeneity within the private sector. Elite schools that levy high annual fees, and that have better overall school characteristics, are significantly better in providing remote instruction and increasing student test scores. This is in line with recent evidence from Andrabi, Bau, Das and Khwaja (2022), who find similar evidence of within-sector heterogeneity in Pakistan, in the context of in-person learning. Furthermore, the results that look at the impact on school quality suggest that private schools are likely to be effective not just during the time of remote instruction, but also in business-as-usual settings, when learning is in person. Given that my data correspond to the period of remote learning, I am unable to test this formally, however prior evidence of private school effectiveness provides findings in support of this, e.g., Muralidharan and Sundararaman (2015); Romero and Singh (2023).

My contributions to the literature are threefold. First, I contribute to the literature on affirmative action in education. There is a large literature on affirmative action that looks into targeting, the mismatch hypothesis, short-term and long-term impacts on the beneficiaries, and cost-benefit analysis, however, most of this work focuses on affirmative action in college admissions.[6] I add to this literature by studying one of the world's largest affirmative action policies that targets children of school entry age, when issues surrounding academic mismatch and fairness in admissions criteria are less of a concern (Romero and Singh, 2023). While this policy has been around for more than a decade, there is very little evidence on its effectiveness, partly because of the recent shift toward centralized admissions, which in turn has facilitated proper record keeping, and access to data. The only other papers that have studied the impact of the RTE quotas on children's outcomes include Damera (2018) and Romero and Singh (2023). I add to this literature by examining a host of mechanisms such as, school quality, parental monetary and time investments, and children's time use (both on the extensive and intensive margin) that might better explain the channels behind gains in children's outcomes. I also contribute by conducting a detailed analysis of the counterfactual destinies of the control compliers, which is useful for the interpretation of the causal effects. Finally I provide evidence from the state of Maharashtra, where the policy implementation rules around the allocation mechanism are very different from the allocation mechanisms implemented in most other states. This is important because the welfare effects of school choice also depend on the allocation mechanism.

Second, I contribute to the extensive literature on school choice, private schools, vouchers, public-private partnerships in education, and education policies in general.[7] In the US, a vast majority of research on school choice focuses on studying the effectiveness of charter schools, which have been found to improve learning outcomes of disadvantaged students, e.g., Cohodes, Setren and Walters (2021). In low- and medium-income countries, the de-

[6] Some examples of this comprise works by Arcidiacono and Lovenheim (2016); Bagde, Epple and Taylor (2016); Bertrand, Hanna and Mullainathan (2010); Bleemer (2022); Card and Krueger (2005); Dillon and Smith (2020), and Khanna (2020).

[7] Glewwe and Muralidharan (2016) provide a review that synthesizes research on education policies combining various developing country contexts.

bate surrounding private schools revolves around concerns of economic stratification and weakening of public schools caused by fee-charging private schools, and potential ways to curtail this, for example, by promoting voucher-like models using public-private partnerships (Glewwe and Muralidharan, 2016). Literature on the relative impact of public and private schools provides mixed evidence.[8] I contribute to this literature by studying lottery-based admissions to private schools through India's RTE policy. My paper provides one of the first estimates from a remote learning context, which offers a unique opportunity to understand private school effectiveness in the context of remote learning, since most of what we know so far about private school effectiveness is from in-person learning contexts.[9] I provide evidence on the distribution of effects within the private sector in the Indian context, and the closest study to do this in a similar context is by Andrabi, Bau, Das and Khwaja (2022) who use value-added models and find substantial heterogeneity within the private and public schooling sectors, in Pakistan.[10]

Third, I contribute to the growing literature on learning loss due to school closures, and ways to mitigate these losses using remote education and technological interventions. A growing number of studies have estimated large learning losses among school children, as a result of the pandemic induced school closures, and recommend post-emergency programs (Azevedo, Hasan, Goldemberg, Geven and Iqbal, 2021; Guariso and Björkman Nyqvist,

[8]With the exception of null impacts of private school vouchers on children's learning in Chile (Hsieh and Urquiola, 2006), most other studies find positive impacts of private schools on learning - PACES program in Columbia (Angrist, Bettinger, Bloom, King and Kremer, 2002; Angrist, Bettinger and Kremer, 2006), private school vouchers in Andhra Pradesh (Muralidharan and Sundararaman, 2015), school value-added in Andhra Pradesh (Singh, 2015). Specifically in the context of India, Muralidharan and Sundararaman (2015) find that private schools achieve these gains at a substantially lower cost per student making them more cost-effective.

[9]A related paper is by Crawfurd, Evans, Hares and Sandefur (2023), who randomize primary school students in Sierra Leone to receive phone tutoring calls from public or private school teachers during the period of COVID-19 school closures. The teachers supplemented government provided radio instruction, but the intervention did not increase children's test scores, whether provided by private or public school teachers. They attribute this non-impact to limited take-up by children.

[10]Prior evidence on heterogeneity within schooling sectors from other country contexts provides mixed evidence - Pop-Eleches and Urquiola (2013) and Jackson (2010) find positive impacts of attending a better school in Romania, and Trindad and Tobago, respectively. In contrast, Abdulkadiroğlu, Angrist and Pathak (2014); Dobbie, Fryer et al. (2011) and Cullen, Jacob and Levitt (2006) find no additional gains on test scores as a result of attending elite and high performing schools in the US.

2023). Another set of studies look at the effect of remote technology interventions on mitigating learning loss (Angrist, Bergman and Matsheng, 2020; Carlana and La Ferrara, 2021; Beam, Mukherjee, Navarro-Sola, Ferdosh and Sarwar, 2021). One such study is by Singh, Romero and Muralidharan (2022), who study a government-run after-school remedial program in Tamil Nadu, India, and find that it was successful in recovering two-thirds of the learning loss in primary school-aged children. I add to this literature by providing evidence of how well-implemented affirmative action policies can act as a safety net for the disadvantaged, during times of severe economic disruptions. In particular, I provide evidence that the policy insured vulnerable children against the risk of non-enrollment, maintained grade progression, and at the same time improved their learning outcomes.

The rest of this paper is structured as follows: Section 2 describes the policy and context (RTE quotas in Maharashtra, and the lottery algorithm); Section 3 describes the data sources (administrative data, and primary data collection) and sampling strategy; Section 4 describes the empirical strategy and also talks about balance, attrition, and external validity; Section 5 discusses results and mechanisms; Section 6 discusses the within-private sector heterogeneity; and Section 7 talks about robustness checks, followed by Appendix tables and figures at the end.

1.3 Background and Policy

The Right to Education (RTE) Act was enacted by the Indian government in 2009, and made education a fundamental right of every child aged 6-14 years. I focus on a specific Clause 12(1)(c) of this act under which all private schools in India are mandated to reserve at least 25% of the seats in entry-level grades for children belonging to low socioeconomic (SES) families.[11] Children who get admitted to private schools under this policy are eligible to get free education from the respective schools until they complete grade 8. The government

[11]Religious and linguistic minority schools are exempted under the RTE Act. Entry level grades comprise grade 1 and pre-primary grades (for example, nursery or kindergarten).

reimburses private schools to cover the school's tuition fee for children admitted under the quota. Children admitted under this quota are also eligible to get free textbooks and uniforms from the respective schools but the enforceability of this varies across states and schools. These quotas were motivated in part due to the rapid increase in fee-charging private schools. Fee-charging private schools accounted for a total of 5.8% of enrollment in rural India in 2002 (Kingdon, 2007), and in more recent years, this has shot up to about 31% primary school enrollment in rural areas, and 50% in urban areas (Pratham, 2019). Due to the rapid growth in demand, there were growing concerns about the rise in segregation within classrooms with the well-off moving to private schools, and the relatively worse-off being in the government schools (which are free of cost). Thus, one of the goals of these quotas is to desegregate classrooms on the basis of socioeconomic status and improving access to quality schooling for all. The quota requirement has been met with restraint across states, and while it was adopted by several states over the years, the policy remains unimplemented in several states (Romero and Singh, 2023).

1.3.1 RTE quotas in Maharashtra: context and lottery mechanism

1.3.1.1 Private school quotas in Maharashtra

I study the impact of this policy in the context of the second most populous state in India, Maharashtra. Maharashtra adopted this policy in 2010 and the eligibility criteria includes children from historically disadvantaged caste groups, low income backgrounds, and children with disabilities.[12] The government reimburses schools for each child who is enrolled under this policy by sponsoring the school fee up to a certain limit and schools are not allowed to

[12] Historically disadvantaged castes include Scheduled Castes, Scheduled Tribes, and Other Backward Classes (OBC). Low income families are defined as those earning less than INR 100,000 per annum ($4746 in PPP). In my administrative data for the year 2020-21, the majority of applications were received under the low income and disadvantaged caste category. Applications received under the disability category comprised 0.6% of the total applications.

charge any fee to the quota students.[13]

1.3.1.2 Online applications

Maharashtra adopted a centralized online application system under this policy, in the academic year of 2017-2018. The online application to apply to schools under this policy begins in the month of February and is open for a month, following which the allocation of students to schools begins based on a centralized lottery algorithm. The majority of schools in the state follow the June to April school year.[14] The process of online application includes filling out the child's details along with household characteristics, for example the child's name, date of birth, gender, and household characteristics like religion, caste and income (if applying under the low income quota). The most important information that is filled out is the house address details, after which the system generates a list of all private schools available under the policy, in the child's neighborhood in three distance bins - all schools available within 1 km radius of the house address, within 1-3 km of the house address, and beyond 3 km of the house address (within the district). This is an important detail of the application process, which I come back to in my estimation strategy. Parents are allowed to choose a maximum of ten schools combining all three distance bins, but they cannot rank schools in order of their preference. They are also required to indicate the eligibility criteria which could be any one of these: low income category, disadvantaged caste category, or child disability category. Finally, parents sign an online declaration which says that in the event of winning a seat, parents are required to show a proof of house address (which must match the address reflected in the online application) and a valid proof that establishes their eligibility criteria under this policy. According to the rules, admission at allotted schools is guaranteed

[13] The reimbursement received by schools is equal to the value determined using the smaller of the these two amounts: school fee charged to fee-paying students, or the upper cap set by the government based on per-pupil expenditure in government schools in the state. The reimbursements have to be borne by centre and state governments in a 60:40 ratio. The policy has been slightly controversial since private schools may choose not to comply with RTE quotas if their fee levels exceed the reimbursement limits. As of year 2020-21, the per child reimbursement under RTE in Maharashtra was capped at INR 17,640 per annum (approximately 213 USD).

[14] A small number of schools follow the May to March school year.

conditional on the house address documentation and other eligibility proofs being valid.[15] Importantly, the declaration states that the documents must be genuine, and in the case that any documents are found to be false or counterfeit, it may lead to monetary penalties and cancellation of the admission offer.[16] Since the policy is targeted towards disadvantaged households, help centers are organized during the weeks of the online application window (oftentimes in schools, and community centers) to specifically assist interested households with filling out the online application and answer questions. Similarly, in the weeks leading up to the start of the online application, the policy is advertised through notifications and billboards outside school premises, community centers, and local newspapers.

1.3.1.3 Lottery algorithm

States have considerable autonomy in how they implement the RTE quotas. Thus, the lottery mechanism that determines the allocation of students to schools under this policy also varies across states. In Maharashtra, it is designed such that each school assigns the highest priority to applicants who reside and applied in the nearest distance bin of the school (within 1 km radius of school, henceforth, distance bin 1), followed by those who reside and applied in the next distance bin (within 1-3 km radius of the school, henceforth, distance bin 2), followed by those who reside and applied in the farthest distance bin (beyond 3 km radius of the school, henceforth, distance bin 3). Hence, the overarching goal is to allocate applicants to schools which are closer to their house address. Importantly, parents are not allowed to submit rank ordered lists and can choose a maximum of ten schools. The lottery mechanism is a two-part process where the first part involves determining applicants who end up winning at a school, and the second part involves determining applicants who end up being waitlisted at a school. Applicants who are neither winners, nor waitlisted by the end, are those who lost at each and every school they applied to. The end result is that each

[15] This could be an income certificate, caste certificate, or disability certificate based on whether the eligibility condition chosen is low income category, disadvantaged caste category, or disability category.

[16] In the administrative data I see that 0.6% of the admission offers were cancelled ex-post, due to false or improper documentation.

applicant has one final lottery outcome which is tied to a unique school - they are either a winner at a unique school, or, waitlisted at a unique school (with a waitlist priority), or, have lost everywhere. In other words, if an applicant is a winner then they only won at one unique school; if they are waitlisted, then they did not win anywhere, but were waitlisted at one unique school; if they are neither a winner, nor waitlisted, then they lost at each and every school they applied to. Appendix Section A.3.1 provides even more detail about the mechanism.

1.3.1.4 RTE School lotteries in Maharashtra, 2020-21

My administrative data corresponds to the universe of applications made under the RTE Act, for private school admissions in the academic year of 2020-21. Private school lotteries in the state were extremely competitive in the 2020-21 academic year. A total of 8848 private schools across the state participated in RTE quota admissions, and received applications from 291,365 children. Of these applicants, 35% won, 39% lost and 26% were waitlisted. Most applications were made under the disadvantaged caste category (63.5%), followed by the low income category. Since the applications under the RTE school lotteries were open only till the end of February 2020, the decision to apply to these school lotteries was made before the COVID-19 pandemic hit India (early March, 2020). However, the decision to take admission (in the event of winning a seat) is likely to have been disrupted due to the nationwide lockdown which was imposed in mid-March and thus unexpectedly coincided with the time when schools were offering admissions.[17]

1.4 Data

My data comes from four sources. First is the administrative data, which gives me details of the universe of children who applied to private school lotteries for grade 1 under the RTE

[17] Because of the nationwide COVID-19 lockdown beginning 24 March 2020, RTE admissions continued to be open through the month of December, 2020. Parents were offered the flexibility to complete the admission formalities either remotely or in-person.

quotas, in the entire state of Maharashtra, for the academic year of 2020-2021. Second is the phone survey data, which I collected during the months of Nov-Dec 2021, by contacting a sample of households who applied to these lotteries (using the phone number provided by the household in their RTE application).[18] Third, I use the U-DISE (Unified District Information System for Education) data which contains the administrative data of school characteristics of the population of schools in India. I use data from the 2019-2020 school year as that contained the most recent information on school characteristics prior to the RTE applications. Finally, I use the administrative data on the annual school fee for all RTE private schools in the state of Maharashtra from the 2019-2020 school year.

1.4.1 Administrative data of RTE quota applications

This data provides the details of the universe of applicants who applied for grade 1 private school lotteries in the state of Maharashtra for the academic year of 2020-2021. These were publicly available at the Maharashtra Education Department website. For each child who applied, there was information about the child's name, child's date of birth, parents' name, parent's contact number, house address, religion, caste, household income, list of private schools chosen by the applicant in the three distance bins (within 1km, 1-3km, beyond 3km), and the distance of each school to the house address of the applicant. For each child who applied, there was detailed information about their lottery outcome and how it evolved over time. To be precise, for every child who applied, there was data on the initial status of the application - whether their application was selected, wait-listed, or not selected anywhere. Each child could only have one of these statuses to begin with.

To explain this in further detail, if a child's application status was declared as selected, then it meant that they had won a seat at one of their preferred schools (if they win, they only win at one school and are excluded from all other schools that they had indicated); if the application status was wait-listed, then it meant that their application was wait-listed

[18] The administrative data on the population of applicants under this policy contained phone numbers of the child's parents which allowed me to conduct phone surveys with applicant households.

at just one of their preferred schools, and they were in the consideration set for admission to this school if a previously selected candidate gave up their seat (each wait-listed child would get a wait-list priority number such that a priority number of 1 would mean that this child would be the next in line for admission, if a vacancy was created at this school. This child was also excluded from all other schools if they had applied to multiple schools); if a child's application was not selected anywhere, then it meant that they were neither selected, nor wait-listed at any school that they had indicated in their application. Over time, the status of the application of a child evolves, and for each selected application, there is data on whether the child formally secured admission to the private school that was allotted to them and the corresponding date on which admission was secured (some students forgo their admissions and this creates vacancies for wait-listed children); for each wait-listed child, whether this child was finally admitted to the school that wait-listed them and if so, when they secured admission.[19]

1.4.2 Primary survey data collection

I conducted phone surveys with a sample of applicants during the months of Nov-Dec 2021, to collect a rich data on children's outcomes, and household characteristics. A total of 4259 applicant households were contacted during this period, and successful interviews were completed with 2329 households (response rate of 55%). For each successful interview attempt, I also conducted a short interview with the applicant child to collect data on their learning outcomes in English and Math.[20] Among the full sample, a total of 695 households provide data on children's learning outcomes.[21] Response rates among winning and non-winning

[19]The status would evolve over time and the website put a notice of the deadlines by which selected candidates must approach their allotted schools to secure admission after which their admission would be null and void. Similar notices were put for the waitlisted candidates along with their priority numbers, and the process would continue to extend admission to candidates with lower priorities, until all seats were filled. Candidates were also sent SMS notifications about the deadlines on their registered contact numbers.

[20]The questions to test children on phone-based assessments come from Romero and Singh (2023) and Angrist, Bergman and Matsheng (2020) and are designed to capture foundational language and numeracy skills. The exact questions administered to children are shown in Figure A.3 in the Appendix.

[21]To minimize non-response bias, the following rule was followed for calling households - each household was attempted to be called up to five times before discarding that number. The protocol was to attempt to

applicant households were about 57.7% and 52.4%, respectively. I discuss attrition and non-response bias in Section 1.5.2 and find that my results are robust to differential attrition, using inverse probability reweighting.

1.4.2.1 Sampling strategy

To select the sample of applicant households for conducting phone-surveys, I design a sampling strategy. It is carefully designed to select a sample of comparable winners and losers under the policy, who are otherwise ex-ante similar in their household location and the school preferences that they listed in the RTE application.

The ideal comparison would involve comparing winners and losers who had the same school preferences by each distance bin, to begin with (as indicated at the time of submitting the online application). However, full stratification of applicants based on their distance bin-specific school preferences eliminates many schools and students from consideration (Abdulkadiroğlu, Angrist, Narita and Pathak, 2017).[22] In order to remedy this, I pick my sample such that the applicants who win and lose the private school lottery are comparable to one another to the extent that they made the same school choices in the *nearest distance bin*, i.e., schools chosen within 1 km radius of the house address; or, in other words, had chosen the same *school vector* in the nearest distance bin.[23] This in turn facilitates the comparison of winners and losers under the policy, who were ex-ante similar in their school preferences in the nearest distance bin and resided in the same geographic location. An important point to note is that the sampling strategy is designed to take into account only those schools which were *oversubscribed*, i.e., schools that conducted lotteries to admit applicants, and

call each household once during: the morning, afternoon and evening of a weekday; once on a Saturday, and once on a Sunday.

[22]The most ideal comparison would involve comparing children who differ in their lottery outcome but indicated the same school choice in each of the three distance bins, as this takes care of their endogenous choice of schools, household location, and their ex-ante likelihood of winning entry into schools as determined by the lottery algorithm. However, implementing this is difficult in practice given the high dimensionality of possible school choices over the full population of applicants.

[23]Throughout the paper, I frequently use the term *school vector* to refer to a unique combination of schools chosen in distance bin 1.

those applicants who were subjected to lotteries. This is a limitation in studies that rely on lottery-based designs since oversubscribed schools may differ from undersubscribed schools, which in turn makes it hard to generalize the findings. The strategy is explained in detail in Appendix Section A.3.2, and a schematic flowchart for the same is given by Appendix Figure A.10.

1.4.2.2 Summary statistics

Table A.1 summarizes the characteristics of applicants in the phone survey and also shows the key variables associated with the applicants and their household characteristics. The average applicant is about 7.6 years old at the time of interviews, slightly more likely to be male, and applied to about 5 schools in the RTE application. Some instances of non-enrollment exist in both the academic years, however, there is improvement in enrollment rates in 2021-22, with the easing of pandemic-related restrictions. Conditional on school enrollment, there is variation in the likelihood of schools providing instruction. Several other variables are summarized, such as monetary and time investments in children, their time use, and performance in phone-based assessments; these comprise my outcome variables.

1.4.3 Administrative data of school characteristics

To get at the characteristics of the school being attended by each child in the sample, I use publicly available data on school characteristics from U-DISE for the 2019-2020 school year. This data covers the population of all private and public schools in India and has rich information on schools. I use this data to construct one of my two measures of school quality. I create a school quality index using Principal Components Analysis (PCA) using data on school infrastructure details, digital facilities, teacher quality, and peer composition. I explain this in more detail in Section 1.7.1.

1.4.4 Administrative data of school fees

I use administrative data on school fees for all the RTE private schools in the state which participated in the RTE lotteries in the 2020-21 year. The data comes from the official website of the State Department of Education, Maharashtra and reflects school fees for the 2020-21 year. This data is used in creating the second measure of school quality, where I define schools to be elite based on the annual fee charged. I explain this in more detail in Section 1.7.1.

1.5 Empirical Strategy

Using the administrative data of applicants who applied for private school admissions for grade 1 under the RTE quotas in the academic year of 2020-21, my goal is to estimate the impact of enrolling in a private school as a quota student on children's educational outcomes. The treatment group comprises the beneficiaries under the policy i.e., those who are enrolled as RTE quota students in private schools and the control group comprises non-quota students who may be attending private schools (as fee-paying students), or government schools (free of cost), and those not enrolled anywhere.

There are two endogeneity concerns here, and I address both of them. First, schools selected at the time of submitting the application are endogenous, and second, conditional on winning, the decision to enrol as a quota student is also an endogenous choice. Both these choices might correlate with unobserved household characteristics which might be simultaneously correlated with children's outcomes. I address both these concerns by using a conditional instrumental variables strategy. The idea is that, given the lottery algorithm, conditional on the school choices listed in the application, winning the lottery to a private school is random.[24] While conditioning on the school choices listed in the application solves

[24]This follows from the lottery algorithm which satisfies the Equal Treatment of Equals (ETE) property (Abdulkadiroğlu, Angrist, Narita and Pathak, 2017). ETE is satisfied when students with the same preferences and priorities have the same chance of getting allocated at any given school. If the object of interest is winning a lottery at a school chosen in distance bin 1, then ETE is satisfied each time there is a group

the endogeneity in unobserved preferences for schools, the second endogenity problem is solved by instrumenting quota enrollment with the indicator of winning the lottery, which in turn is random conditional on controlling for the school choices that were listed in the application. Thus, I estimate the local average treatment effect of being enrolled as a quota student on children's outcomes in an instrumental variables framework.

As I explain in the previous section, given the high dimensionality of school preferences, my sampling strategy is designed such that I can condition for the vector of schools chosen in bin 1 and compare applicants who are similar to the extent that they had the same school preferences in bin 1. Conditioning on the vector of schools chosen in bin 1 is one way of addressing the endogeneity in school preferences listed at the time of the application. However, note that given the lottery algorithm and the ETE property, the relevant instrument to be used in such a case is winning the lottery in distance bin 1, which in turn means that the causal effect is estimated for compliers, defined by those who attend private schools as quota students because of winning the lottery in bin 1, and those who don't because they lost lotteries at bin 1 schools. On the other hand, if the instrument is winning the lottery in *any* distance bin, that leads to a much more heterogeneous composition of compliers, i.e., those who are quota students because of winning the lottery in any bin, and those who are not quota students because of losing the lottery in all bins.

Such an estimation can be executed by conditioning on the simulated ex-ante propensity scores of winning the private school lottery (Abdulkadiroğlu, Angrist, Narita and Pathak, 2017). This strategy is useful because it helps reduce the dimensionality of preferences and does not require me to explicitly control for the schools chosen at the time of application. The idea is the following: taking the distance-bin-specific school preferences of applicants as given, one can simulate the lottery algorithm a large number of times to arrive at the

of applicants who had listed the exact same schools in distance bin 1. If the object of interest is winning a lottery at a school chosen in distance bin 2, then ETE is satisfied each time there is a group of applicants who had listed the exact same schools in distance bin 1, and distance bin 2. Finally, if the object of interest is winning a lottery at a school chosen in distance bin 3 or, winning a lottery at *any* school in any of the three distance bins, then ETE is satisfied each time there is a group of applicants who had listed the exact same schools in each of the three distance bins.

simulated ex-ante likelihood of winning the private school lottery, for each applicant. Since the simulated likelihood or propensity score takes into account the school preferences that were listed by the applicant, controlling for these propensity scores essentially performs a similar function as is achieved by explicitly controlling for the full set of schools chosen at the time of application. Since the goal is to estimate the LATE of being enrolled as a quota student under the RTE policy, the identifying assumption in this estimation strategy is that winning the lottery to a private school is conditionally exogenous after controlling for the ex-ante propensity scores of winning the lottery. Below I discuss the implementation of this strategy, which is my preferred approach.[25,26]

Following Abdulkadiroğlu, Angrist, Narita and Pathak (2017), my preferred estimation strategy involves controlling for the vector of dummies of narrow bins of ex-ante propensity scores of winning a lottery in any distance bin.[27] This strategy relies on comparing the winners and losers of private school lotteries, who had a similar ex-ante propensity of winning the lottery to an RTE private school (in any distance bin). This exploits the within-variation that results from comparing winners and losers who had a similar ex-ante propensity of winning any private school lottery, and does not require them to have chosen the same sets of schools. I estimate this using a two-stage least squares (2SLS) procedure, where the first stage is the effect of a random assignment of a private school seat on enrollment, and the second stage estimates the impact of quota enrollment on student outcomes.

[25] I discuss the calculation of these propensity scores in Appendix Section A.3.3. I also show the distribution of these ex-ante propensity scores (Appendix Figure B.1). Appendix Table A.22 shows the detailed distribution of simulated propensity scores for the full population, and the sample.

[26] This strategy is powerful to deal with issues of stratification and sampling such as the one caused by fully stratifying applicants on the basis of their distance-bin-specific school preferences. It relies on comparing winners and losers who had a similar ex-ante likelihood of winning and does not require them to have chosen the exact same set of schools, thus bypassing some of the power issues which may occur if comparisons are based on controlling for school fixed effects.

[27] In the Appendix, I present results that condition on the school vector chosen in bin 1, and compare these results to the case which conditions on the simulated ex-ante propensity of winning in bin 1. Tables A.9, A.10, A.11 show the results for the main outcomes. The two specifications produce very similar results thus providing confidence in the fact that conditioning on simulated ex-ante propensity scores performs a similar function as is achieved by conditioning on the school vectors.

I estimate the following equations via 2SLS:

$$RTE\_Enrolled_i = \alpha_1 WinningLotteryAnyBin_i + X_i'\alpha_2 + \sum_{x=1}^{100} \gamma_x d_i(x) + \epsilon_i \tag{1.1}$$

$$Y_i = \beta_1 \widehat{RTE\_Enrolled}_i + X_i'\beta_2 + \sum_{x=1}^{100} \delta_x d_i(x) + e_i \tag{1.2}$$

where, $d_i(x)$ are dummies taking a value of 1 if child i's estimated propensity score of winning a lottery at a private school in any bin lies in the respective 0.01 wide probability bin, X_i is the vector of child and household characteristics like sex and age of child, indicator for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. These covariates are added only to increase the precision of my estimates and the results are robust to excluding them. The coefficient of interest is given by β_1, which captures the LATE of attending a private school as a quota student on child outcomes. The compliers are those who attend private schools as quota students because they won the lottery to a private school (in any bin), and those who are without a quota because they lost the lottery at all schools that they listed in their application, and may be attending private schools as fee-paying students, or government schools, or may be out-of-school.

For some of the surveyed households, responses on certain conditioning variables are sometimes missing. Instead of a listwise deletion of observations that have missing values for covariates, I re-code missing values of covariates to their mean value in the sample and control for these re-coded covariates, and include a separate missing value indicator in all the specifications. Listwise deletion of observations missing any of the conditioning variables would mean non-randomly dropping a substantial fraction of the sample (King, Honaker, Joseph and Scheve, 2001; Black, Smith and Daniel, 2005).

1.5.1 Balance

I test for balance across winning and non-winning applicants to examine if they are similar on baseline observed characteristics. Table A.2 presents the results, conditioning on the ex-ante propensity of winning at any bin. The majority of the characteristics are balanced across the two groups, with some exceptions - for example, father's education, religion, and household SES index. This suggests that the winners and non-winners are modestly balanced on baseline observed characteristics.

1.5.2 Attrition

A concern that could potentially bias estimates is whether there is selection into who agrees to be a part of the phone surveys. For example, if winners were more likely to participate in the survey, and at the same time also benefited from the quota seat, then this could bias the effect sizes in the upward direction. Table A.3 shows whether there is selection into participation in phone surveys based on observable characteristics of households at baseline, after conditioning on the ex-ante propensity of winning in any bin.[28] The table shows this for household's participation in phone surveys, and for household's participation in phone-based assessments with the applicant child, conditional on being part of the phone surveys. As can be seen from Panel A, attrition is slightly unbalanced - winning applicants were 5.7pp more likely to agree to be interviewed relative to the non-winning applicants. However, there is no systematic attrition by winning status, on participation in child assessments, conditional on survey participation. I test for robustness of my results on the main outcome (phone-based assessments), using inverse-probability reweighting to account for differential attrition (Table A.19).

1.5.3 External validity

My results are based on a lottery-based research design. While lottery-based estimates help in removing selection bias, there are several challenges with this design. First, these estimates are specific to oversubscribed schools, which might be different from undersubscribed schools.

[28] Results are robust to conditioning on school vector fixed effects of winning in bin 1 or the ex-ante propensity of winning in bin 1.

For example, oversubscribed schools might be overrepresentatitve of urban areas, relative to rural areas.[29] Second, it relies on applicants who faced lotteries to get admitted to schools, a group that may differ from nonapplicants (Angrist, Cohodes, Dynarski, Pathak and Walters, 2016). Third, the LATE identifies a treatment effect only for compliers which is a very specific sub-population of the treated (Black, Joo, LaLonde, Smith and Taylor, 2022). Nevertheless, Kline, Rose and Walters (2022) show that LATE is the policy-relevant parameter in case of a marginal increase in the number of available seats among lottery applicants (Angrist, Hull and Walters, 2023). I discuss the issue of external validity in more detail in Section 1.5.3.1, where I discuss complier characteristics - these can provide a partial guide to external validity in the context of lottery-based IV estimates (Angrist, Hull and Walters, 2023).

1.5.3.1 Characterizing Compliers

The instrumental variables strategy identifies a unique causal parameter, which is specific to the sub-population of compliers for that instrument. Different valid instruments for the same causal relation therefore estimate different things, because the compliers are essentially different based on the instrument (Angrist and Pischke, 2009). Since the IV identifies the average treatment effect for the compliers, it is a useful exercise to learn more about the characteristics of the compliers. Another important reason to study complier characteristics is that they can provide insights about external validity of a set of lottery-based IV estimates (Angrist, Hull and Walters, 2023).

I use Angrist, Hull and Walters (2023)'s implementation of the methods discussed in Abadie (2002), to compute complier characteristics. Table A.4 shows the differences in baseline characteristics of the compliers, always- and never-takers in Maharashtra's RTE lottery. The table shows the mean of baseline characteristics for each of these groups (see

[29] Romero and Singh (2023) compare the lottery-based estimates to a random sample of applicants who are always assigned to a private school and find that the lottery-based sample of students is moderately better off than the sample of students with a guaranteed private school allocation. They point out that this might be a function of urban areas being over-represented in their core sample, which have more oversubscribed schools. This has also been observed in charter school lotteries in the US (Cohodes, Setren and Walters, 2021).

Appendix Section A.4.1 for details on implementation). Untreated and treated compliers are very similar across all characteristics as shown in columns (1) and (2). Columns (3) and (4) show the mean characteristics for always- and never-takers. Relative to all other groups, always-takers are slightly more likely to be low income quota applicants, Muslims, and households with mothers having finished primary education. Relative to the other two groups, the average complier is slighlty more likely to be Hindu, and less likely to be from Scheduled castes. However the magnitude of the differences are small indicating that overall group characteristics are quite similar across groups. Overall, this suggests that compliers are representative of the full sample of applicants and that the external validity of the LATE extends to always- and never-takers.

1.6 Results

1.6.1 First stage

Table 1.1 shows the first stage which captures the relationship between lottery offers and enrollment as a quota student. The endogenous variable of interest, i.e., enrollment in a private school as a quota student is instrumented by the indicator of winning the lottery at a private school, under the RTE policy. The instrument is random conditional on controlling for the narrow bins of simulated ex-ante propensity scores of winning. The results show that winning the lottery is strongly and positively correlated with enrollment as a quota student. The first-stage estimates are smaller than one because of non-compliance among lottery winners - some lottery winners choose to opt out of the quota seat at allotted schools as they may prefer other schools.[30] Another reason for the reduced estimate of the first stage is that some applicants who did not win any lotteries in the beginning, received an

[30]Some of this non-compliance may also stem from the fact that the timing of seeking admissions at allotted schools under the RTE policy coincided with the COVID-19 lockdown. However, the extent of COVID-19 induced non-compliance among lottery winners was reduced to some extent, as a result of schools allowing admission formalities to be completed over the phone.

offer through the waitlist (at a later date).[31]

Table 1.1: First stage of winning the RTE lottery on enrollment as a RTE quota student

	Enrolled as RTE student
	(1)
Instrument = Winning the Lottery (any bin)	0.792***
	(0.013)
Outcome mean	0.44
Control mean	0.08
Observations	2,329
R^2	0.66
Pscores of winning	Yes
Controls	Yes

Notes: This table shows the first stage effects of winning the RTE private school lottery in any distance bin on enrollment as an RTE quota student in a private school. This first stage corresponds to the 2SLS regression where the outcome of interest is school enrollment. Control variables include sex and age of child, dummy of father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

1.6.2 Primary outcomes

There are two primary categories of outcomes of interest. First, enrollment in the two academic years and second, performance on phone based assessments. Table 1.2 shows statistically significant gains in enrollment in both academic years for treated compliers. Gains in enrollment are approximately three times higher in 2020-21 relative to 2021-22 suggesting that some of the children who were out of school during the first year of the pandemic (in 2020-21), are enrolled in schools in the following year (2021-22), with the easing in restrictions surrounding the pandemic. Column (3) sheds light on another important aspect which is that RTE quota students were more likely to be able to maintain the right grade-for-age trajectory following their timely enrollment.

Being a quota student not only increases the likelihood of being enrolled and being enrolled in a higher grade, but also leads to gains in performance on phone-based assessments. As can be seen in Table 1.3, there is a 0.18 SD unit increase in English performance for the

[31]The first stage estimates differ across outcomes due to changes in sample composition - for example, the phone-based assessments are for a sub-sample of the surveyed households and some outcomes have missing responses leading to a reduced sample.

Table 1.2: LATE of being a RTE quota student on enrollment

	Enrollment (2020-21)	Enrollment (2021-22)	Grade 2 and above (2021-22)
	(1)	(2)	(3)
Enrolled as RTE student	0.141***	0.048***	0.194***
	(0.016)	(0.009)	(0.017)
First stage F-stat	3,911.06	3,938.19	3,934.19
Outcome mean	0.89	0.97	0.86
Control mean	0.84	0.94	0.78
Observations	2,328	2,328	2,327
R^2	0.10	0.07	0.15
Pscores of winning	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on children's enrollment. The outcomes in columns (1) and (2) measure the indicator of school enrollment in the two academic years. Column (3) measures the indicator for whether the child is in grade 2 or grade 3 in the 2021-22 academic year. Controls include sex and age of child, dummy of father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

treated compliers. Although, the effect on Math is statistically indistinguishable from zero (at conventional levels), it is quite similar in magnitude to English.

Table 1.3: LATE of being a RTE quota student on phone based assessments

	Test score (standardized)	
	English	Math
	(1)	(2)
Enrolled as RTE student	0.187**	0.144
	(0.089)	(0.091)
First stage F-stat	1,129.88	1,129.88
Outcome mean	-0.00	-0.00
Control mean	-0.10	-0.09
Observations	695	695
R^2	0.17	0.13
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on children's performance on phone-based assessments. Outcomes measure children's standardized test scores on English and Math. Controls include sex and age of child, dummy of father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Another thing to pay attention to, is to understand the composition of compliers, since the causal impact of interest is relevant to this group. In particular, while the treated compliers comprise a homogeneous group of students (enrolled as quota students because of winning the lottery), the same is not true for the control compliers. The latter group comprises fee-paying students at private schools, students who go to government schools, and students who are out of school. Thus, it is helpful to characterize the distribution of enrollment status across these various sectors (Angrist, Hull and Walters, 2023). Abdulkadiroğlu, Angrist and Pathak (2014) and Chabrier, Cohodes and Oreopoulos (2016) refer to this as *counterfactual destinies*. Table A.5 shows the counterfactual destinies for control compliers - 65% of the lottery losers end up enrolling in private schools as fee-paying students, 20% in government schools and about 5% are out of school.[32] Aside from the caveat that the control compliers are not a homogeneous group, and that these gains in test scores reflect differences in the mean outcomes of treated compliers relative to control compliers (who might vary in their enrollment status on the extensive margin, and type of school in the intensive margin), the effect sizes are similar to those observed in Muralidharan and Sundararaman (2015) who estimate the intent-to-treat (ITT) effects of being awarded vouchers to private schools. After 4 years of program implementation, they find that winners of private school vouchers perform 0.12 SD units better on English in the state of Andhra Pradesh, India. Similar to my findings, they find null impacts on Math. Their mechanisms show that these effects are primarily driven by the differences in instructional time spent across subjects.

In the next section, I investigate multiple mechanisms that might cause these gains. One mechanism that stands out and might explain gains in English, is the increase in likelihood of attending English medium schools - Table 1.4 shows that treated compliers are approximately 9 percentage points more likely to be in English medium schools relative to control compliers.[33] However, other channels, such as, instructional time spent across subjects, and quality of instruction, may also play a role in explaining these effects. While my survey

[32] Among lottery losers, there are some children for whom the school name and the official school code could not be matched with the administrative data on the population of schools. Thus, for these children, the school sector – private, government, or out of school – is missing. It is for this reason that the counterfactual destinies don't add up to one.

[33] In India, English medium schools refer to schools where the primary language of classroom instruction is English. English-medium instruction is also perceived to have large labor market returns, see, e.g., Azam, Chin and Prakash (2013).

does not collect data on instructional time per subject, it contains other rich information on school-specific instruction that I talk about in the next section.

The majority of existing evidence of private school effectiveness is in the context of in-person learning. More recently, since the COVID-19 pandemic, there has been a growing interest in understanding the impact of remote instruction on children's educational outcomes, but we know relatively little about how schools adapt to changes in learning environments, and how this varies by school sectors, and whether private schools are still differentially effective in the context of remote learning.[34] If private schools are seen to be effective in remote learning environments, then this might be especially relevant in developing country contexts where private school penetration is low or skewed. While constructing schools to provide uniform access to quality schooling might be the long term goal of governments, a short term cost-effective solution could be to increase access to private schools through remote learning. My results suggest that virtual learning can be effective, and that private schools do a better job at adapting to, and implementing, remote educational technologies, and in doing so, they also enhance children's learning.

1.6.3 Mechanisms

Children's cognitive achievement and human development are considered to be a cumulative process that depends on the history of family and school inputs, and on children's innate ability (Becker and Tomes, 1976; Todd and Wolpin, 2003). Following that, I explore the various mechanisms that might explain these improvements in test scores. In particular I study three channels - school inputs, parental inputs, and children's own time use and educational effort.

[34] A related paper is by Crawfurd, Evans, Hares and Sandefur (2023), who randomize primary school students in Sierra Leone to receive phone tutoring calls from public or private school teachers during the period of COVID-19 school closures. The teachers supplemented government provided radio instruction, but the intervention did not increase children's test scores, whether provided by private or public school teachers. They attribute this non-impact to limited take-up by children.

1.6.3.1 School inputs

First, I discuss the channel of school inputs and school quality. Table 1.4 looks at school characteristics that might matter in children's educational production function. The first two columns look into the outcomes of attending a private school and whether the school's primary language of instruction is English. Being enrolled as an RTE quota student increases the likelihood of both these outcomes, for the compliers. A notable observation is the magnitude of these effects - the likelihood of attending a private school increases only by 20 percentage points. However, it is not surprising to see a small effect size, given the evidence of gradual exodus of children from government schools as a result of the increased affordability and demand for private schools (Kingdon, 2020). In the control group (non-quota students), about 79% of the children are enrolled in private schools, suggesting that for many applicant households, the RTE policy might just be a way of upgrading to a *better* or a more *expensive* school within the private sector - a finding that Romero and Singh (2023) investigate in greater detail in the context of RTE lotteries in the state of Chhattisgarh, India. This also points to the fact that there is substantial variation in the quality of schools *within* the private sector, and that this might matter in determining children's educational outcomes; I investigate this point in further detail in Section 1.7.

Table 1.4 also shows that conditional on being enrolled, treated compliers were more likely to be enrolled in schools that were actively providing instruction in the two academic years (columns (3) and (4)). The magnitude of the effect size is larger in the 2020-21 academic year (7 percentage points), relative to the 2021-22 academic year (3 percentage points). This indicates that RTE schools were especially more effective in providing remote instruction during the year that coincided with pandemic-induced school closures, and that this effect size was reduced by about half in the following academic year. The reduction might come from a combination of these channels - schools being attended by control compliers might be getting better at providing remote instruction over time, and the extent of being out of school is falling among control compliers.

Table 1.4: LATE of being a RTE quota student on school inputs

	School type		School instruction		School instruction modality		
	Private (2021-22)	English medium (2021-22)	Provides instruction (2020-21)	Provides instruction (2021-22)	Synchronous (online) (2021-22)	Recordings shared (audio/video) (2021-22)	Activity plans (WhatsApp/SMS) (2021-22)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Enrolled as RTE student	0.199***	0.089***	0.072***	0.030***	0.136***	-0.034*	-0.066**
	(0.016)	(0.013)	(0.017)	(0.007)	(0.021)	(0.018)	(0.026)
First stage F-stat	3,856.22	3,859.05	3,472.00	3,877.71	3,788.83	3,788.83	3,788.83
Outcome mean	0.88	0.94	0.89	0.98	0.77	0.14	0.55
Control mean	0.79	0.89	0.85	0.97	0.70	0.15	0.57
Observations	2,249	2,250	2,083	2,255	2,210	2,210	2,210
R^2	0.16	0.08	0.11	0.05	0.15	0.08	0.10
Pscores of winning	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on school inputs. Columns (1) and (2) show the indicator of a child being in a private school, and the school having English as the primary language of instruction. Columns (3) and (4) show outcomes for whether the school provides instruction in the two academic years. Columns (5) - (7) show the type of instruction modality offered at the child's school in the 2021-22 academic year. The question was: in the past month, what were the types of instruction offered by child's school (select all that apply) - (i) online classes with teacher, other students (ii) pre-recorded lectures were sent (audio/video) (iii) written learning activity plans were shared via Whatsapp/SMS (iv) other, specify. This question was asked only to children who were enrolled in school in 2021-22, and whose school was providing instruction. Controls include sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Finally, in the last three columns I look at the modality of instruction being offered at the school, conditional on being enrolled. Treated compliers are 13.6 percentage points more likely to be enrolled in schools that offer synchronous (online) classes, whereas control compliers are more likely to be in schools that offer recordings of lectures and share text-based activity plans (via WhatsApp or SMS).

Another outcome of interest is how winning the school lottery improves the overall quality of the school that a child attends. This is of specific interest in the context of school choice in the US where school quality is often defined using peer achievement and/or peer socioeconomic composition etc. Appendix Table A.6 shows the impact on being a quota student on school quality, conditional on enrollment. In the absence of data on peer achievement, I measure school quality by creating indices of specific broad categories of school-level characteristics using principal component analysis (PCA). I find that quota students enroll in schools that have better infrastructure facilities, digital facilities, teacher quality, and have a less diverse student composition. Interestingly, comparing the magnitudes across specific indices suggests that digital and teacher indices increase more than infrastructure, suggesting that gains in performance of children are likely to be driven more because of school characteristics that actually matter for remote instruction, rather than infrastructure facilities which are less likely to directly matter for remote instruction.

Finally, Table A.7 shows the LATE of winning the lottery on the likelihood that the school the student attends teaches any given subject, after conditioning on enrollment. Relative to the non-quota students, quota students are more likely to be in schools that teach English, Hindi, Environmental studies, Computers, General knowledge, Arts, Music and Dance. However, they are no more likely to teach Math. A caveat is that this information comes from parental responses, and not from the school, so there could be measurement error in the data. But with that caveat aside, these estimates suggest that gains in English could be driven through a combination of reasons. Firstly, schools are more likely to teach English. Secondly they teach more subjects, and since the primary language of instruction is

English it is likely to further complement students' understanding of the English language.

Overall, these results suggest that treated compliers end up in schools that were more likely to provide instruction and also provide synchronous modes of instruction - which is arguably more effective and holds both teachers and students more accountable, by offering real-time interaction. Furthermore, these schools are better equipped with digital facilities and have teachers with more qualifications, both of which might matter for augmenting children's learning during the period of remote instruction.

1.6.3.2 Parental inputs

Second, I explore the channel of parental investments in children as receipt of a quota seat may also change household inputs into education (Das et al., 2013; Pop-Eleches and Urquiola, 2013). If parental investments change as a result of quota receipt, then the LATE of attending a private school as a quota student on test scores reflects an overall effect of school inputs and home inputs on children's achievement (Becker and Tomes, 1976; Todd and Wolpin, 2003). Recent research in this literature attempts to understand whether public investments in children encourage or crowd out parental investments - knowledge of this can inform policy and improve the targeting of public funds towards school inputs that encourage parental effort (Rabe, 2020). I test the extent to which parents adjust their time and monetary investments in children in response to winning the quota seat. I find that parents increase their investments in winners, however the effect sizes are small, indicating that even though parental inputs are increasing, they only explain a small part of the story.

I collect detailed data on parental monetary and time investments which help me in studying both the extensive and intensive margin impacts. Table 1.5 shows that while there is no statistically significant impact on the extensive margin of children receiving household help with educational activities, there is evidence on the intensive margin as treated compliers are likely to receive a little more help per week with educational activities (approximately 30 mins more per week), relative to the control compliers. Further analyzing this, I find that

Table 1.5: LATE of being a RTE quota student on parental investments

	Time investments		Monetary investments	
	Receives help with homework	Hours of help (hrs/week)	Any expense (past year)	Expenditure (past year)
	(1)	(2)	(3)	(4)
Enrolled as RTE student	0.020	0.543*	0.065***	-90.182
	(0.013)	(0.308)	(0.014)	(158.303)
First stage F-stat	3,915.91	3,915.91	3,822.52	3,822.52
Outcome mean	0.93	9.50	0.93	3,462.86
Control mean	0.92	9.31	0.91	3,467.37
Observations	2,329	2,329	2,227	2,227
R^2	0.08	0.06	0.06	0.06
Pscores of winning	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on parental time and monetary investments in children - both on the extensive and intensive margins. Column (1) measures the extensive margin of whether the child receives any help with educational activities in the household, and column (2) measures the intensive margin of the number of hours of help. The survey questions were: "Does the child receive any help with educational activities from any members of the household?" followed by details of each person who helps and their relationship with the child. Next, it was asked: "Among all those who help, who is the person who most often helps the child with educational activities?", followed by details about number of hours per day of help on a typical day, and number of days of help per week in the past week, to calculate weekly hours of help coming from the main helper. Hence, data on hours of help are collected only for the main helper. Column (3) measures the extensive margin of any educational expenses in the child in the *past one year* (on curriculum books, notebooks, and stationary), and column (4) measures the intensive margin of the amount of expenditure incurred on child's education in the past one year. There are some missing values for the monetary investment questions due to item non-response. Controls include sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

among all the household members, mothers are the ones who are statistically significantly more likely to help children with educational activities (Table A.8). Turning to monetary investments, I find that parents of lottery winners are 6 percentage points more likely to spend on the educational needs of children in the past year (on curriculum books, and stationary). However, there is no detectable impact on the intensive margin.

Together, these results suggest that parents respond to the receipt of the RTE quota seat by reinforcing investments in children, however the effect sizes suggest that this channel represents only a small part of the story.

1.6.3.3 Children's time use

Third, I explore the channel of children's own effort by looking at children's time use. Table 1.6 shows the LATE of being a quota student on children's time use in educational activities, and I find that being a quota student statistically significantly increases children's time use in educational activities. Treated compliers spend 3 more hours per week doing school-related activities, and approximately 20 more minutes per week doing homework. I don't find any statistically significant differences in time spent on private tutoring (after school classes) and in non-educational activities. Overall, children seem to increase their educational effort as a result of winning the lottery and this might also contribute to gains in their educational outcomes. These findings are also consistent with Muralidharan and Sundararaman (2015), who find no impact of winning private school vouchers on home study- and play-habits except increased time spent in school for voucher winners.

Table 1.6: LATE of being a RTE quota student on children's time use

	Educational activities			Non-educational activities		
	School (hrs/week)	Tuition (hrs/week)	Homework (hrs/day)	Playing (hrs/day)	Television (hrs/day)	House chores (hrs/day)
	(1)	(2)	(3)	(4)	(5)	(6)
Enrolled as RTE student	2.945***	-0.405	0.262***	-0.099	-0.073	-0.006
	(0.400)	(0.315)	(0.038)	(0.062)	(0.047)	(0.021)
First stage F-stat	3,915.91	3,915.91	3,915.91	3,909.99	3,938.97	3,915.91
Outcome mean	12.12	4.67	1.40	2.45	1.10	0.39
Control mean	10.79	4.92	1.31	2.50	1.15	0.40
Observations	2,329	2,329	2,329	2,328	2,322	2,329
R^2	0.13	0.08	0.06	0.05	0.09	0.06
Pscores of winning	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on children's time use in educational and non-educational activities. School hours are set equal to zero for those who report being not enrolled in any school. Tutoring hours (differs from formal schooling, typically happens after school) are also set equal to zero for those who report being not enrolled in any private tuition. The question for homework hours is not always zero for not enrolled children, as the question asked - "how much time does child spend doing homework, or any educational activities after school?". There are some missing values for playing and watching television due to item non-response. All time use data are winsorized at the 99th percentile. Controls include sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

1.7 Winning in Elite versus Budget RTE private schools

The results so far provide evidence that winning entry to RTE private schools statistically significantly improves children's learning outcomes. The mechanism analysis suggests that school's mode of instruction, and children's effort in educational activities play important roles in achieving these gains. However, even among the class of RTE private schools attended by winners, there might be variation in school quality that makes some private schools better relative to others. Private schools that levy a high yearly school-fee are likely to have highly qualified and motivated teachers with high teacher salaries, and thus offer higher quality of education, better resources, and as a result might have a higher value-added.[35] In contrast, private schools that charge a lower fee might have fewer teachers (and thus larger class sizes) with fewer qualifications, and as a result have lower value-added. Thus, even among the RTE winners who benefit from a quota seat at private schools, school quality is likely to differ, which may lead to differences in children's achievement. But do these differences in school quality matter during periods of remote learning?

There is no evidence on the how the distribution of school effectiveness varies within the private sector in India (Romero and Singh, 2023). The only such evidence from a similar context is from Punjab in Pakistan, by Andrabi et al. (2022). Using value-added models (VAMs), they find evidence of substantial within-village variation in school quality within the private schooling sector. Contrary to the existing literature that has largely focused on a single private school premium, their findings suggest a range of causal estimates of the private school premium, resulting from a substantial *within-sector* variation in school quality. If there exists variation in quality within private schools, then it might be misleading to focus

[35] Previous literature has used school fees as a proxy for school quality. Rao (2019) defines elite schools as those charging a fee greater than 2000 INR per month, in New Delhi. Andrabi et al. (2022) find a positive correlation between school value-added (SVA) and school fees in Pakistan. Romero and Singh (2023) analyze the impact of winning a quota seat under RTE on the market price of the school being attended, in the state of Chhattisgarh, India, and find that quota students enroll in costlier schools.

on a single estimate of the private school premium. The next question that arises is how one should arrive at a reliable measure of school quality. In this section, I uncover the impacts of relative differences in private school quality on children's educational outcomes by using two alternative measures of school quality which I discuss in detail in Section 1.7.1. The next paragraph briefly summarizes how the literature defines school quality.

The idea of school quality is a latent concept, and the literature has looked at various ways of measuring the *true* school quality. The bulk of the literature on school quality in the US focuses on achievement-based measures of quality, and more recently on outcomes other than student achievement, for example, crime, employment, earnings and non-cognitive outcomes (Angrist, Hull and Walters, 2023).[36] Several papers use peer ability and socioeconomic composition of peers as proxy for school quality (Abdulkadiroğlu, Angrist and Pathak, 2014; Pop-Eleches and Urquiola, 2013; Dobbie and Fryer Jr, 2013). Greaves et al. (2023) use school inspection ratings as a source of information on school quality in the context of England. School management interventions that improve the quality of leadership practices have also been utilized to get at measures of school quality.[37] Specific dimensions of school inputs have also been used to measure school quality, such as class size (Datar and Mason, 2008; Fredriksson, Öckert and Oosterbeek, 2016) or school resources (Houtenville and Conway, 2008; Das et al., 2013). Another important and related strand of literature is on college quality, where the goal is to study the educational and labor market effects of the quality of college that individuals attend. Black and Smith (2006) discuss the issues with using a single proxy of college quality (such as the average SAT score of the entering class) as it leads to substantial measurement error in the quality measure, and propose several solutions. One of the proposed solutions is to create a quality index that combines multiple individual quality

[36] Angrist, Hull and Walters (2023) provide a useful review of this literature by summarizing the various econometric strategies for estimating school effectiveness - school lotteries using the instrumental variables approach, regression-discontinuity approach where students are admitted based on a cutoff score, centralized school assignment where school allotment happens via conditional randomization based on rank ordered lists submitted by parents, and finally value-added models (VAMs) which control for lagged outcomes and covariates by making use of panel data of student test scores.

[37] Anand et al. (2023) conduct a meta analysis of the impact of school management interventions on student learning using data from multiple evaluations, and provide a systematic review of this literature.

measures (or proxies) via factor analysis. The larger the number of quality variables, the less is the measurement error in the index (Black, Smith and Daniel, 2005). I take inspiration from them to define one of my school quality measures in a similar way, as I discuss in more detail in the next section.

1.7.1 Two alternate measures of school quality and eliteness

In the absence of panel data on standardized test scores across schools, and school level peer achievement, I consider two ways of defining school quality - one that uses school fees, and another that uses data on a rich set of school-level characteristics.[38] I start with a simple case of categorizing schools as *elite* or *budget*, based on two measures of quality.[39] The first measure uses administrative data of schools' annual fee to categorize each school as elite or budget. Figure A.4 shows the distribution of annual school fee for private schools in the state using the administrative data of annual school fee charged by RTE private schools in the state. As the figure shows, most of the private schools are concentrated on the lower end of the fee distribution. This is in line with Kingdon (2020), who documents that the vast bulk of private schools in India are low-fee schools, when benchmarked against the state per capita income and daily wage laborer's incomes. The author also points out that this increase in affordability has led to a rapid migration of students towards private schools, and an emptying of government schools. Taking the distribution of annual school fee for all the RTE private schools in the state, I define a school as elite if the annual school fees exceeds

[38]The school fees based measure is in the spirit of previous literature that uses fees to define school eliteness and quality, such as Rao (2019); Romero and Singh (2023) and Andrabi, Bau, Das and Khwaja (2022). The school-quality-based measure is in the spirit of the college quality literature that uses multiple college characteristics to create an index of college quality, such as Black, Smith and Daniel (2005) and Black and Smith (2006).

[39]Note that, while both my measures of quality are continuous measures of quality, the discretization of schools into elite and budget is done following the identification strategy which relies on the within-variation in lottery outcomes of children with similar ex-ante propensities of winning the lottery *at elite* schools. This in turn requires that each school that was chosen during the time of application, be categorized as a binary of either elite or not-elite to get at the simulated ex-ante propensity score of winning at an elite school. Black and Smith (2006) discuss that such discretization leads to a loss of information and in turn causes researcher-induced measurement error in the quality index.

the 75^{th} percentile in the distribution of fees of all private schools in the state, and budget, otherwise.[40]

The second measure of school eliteness is based on a school quality index that I construct using Principal Component Analysis (PCA). The data for this analysis comes from UDISE, which allows me to make use of a rich dataset on school characteristics – infrastructure details, digital facilities, teacher qualifications, and peer SES composition – which might matter in determining the overall school quality. The complete list of variables that are used for creating this index is shown in Figure A.5 in the Appendix. I use the first component of the PCA to create the quality index. The figure also shows the factor loadings on each of the variable - it shows that all these different types of school inputs are positively associated with school quality.

These two measures of school quality display a strong and positive correlation - Table A.13 shows the results from OLS estimates of a simple linear model where I regress school fees on school's PCA based quality index.[41] While both measures capture school quality, I prefer the fee-based measure to the PCA index. The reason for this is that school fee is likely to encapsulate various dimensions of school quality - both observed and unobserved - which may not be apparent in the school characteristics data that are used to construct the PCA index. The school characteristics data is useful to the extent that it provides information about observed characteristics of the school. The fee on the other hand is a measure that is likely to take into account all the aspects about schools which may be hard to quantify.[42]

[40] I vary the bar of eliteness by lowering and increasing the threshold to the 50^{th} and 90^{th} percentiles, respectively. In the sub-sample of lottery winners, only 2% of the children attend elite schools but not as an RTE student. In the sub-sample of lottery winners (which is the relevant sample for this exercise), approximately 4% of the children attend government schools. I assume that government schools (free of cost) are budget schools. Most of the government schools being attended by the non-compliers among the lottery winners are zilla parishad schools (state-run schools which are established, funded and supervised by the district councils of India), which lie at the lower end of government school quality distribution.

[41] Tabulating schools on these two measures of eliteness shows that majority of the schools are consistent in the elite definition across the two measures (Figure A.6). About 65% of the schools are elite on both measures, about 25% of the schools differ in classification of eliteness across the two measures, the rest have missing data on one of the two measures.

[42] Note that using a fee-based measure as a quality measure assumes that schools don't increase fees in response to the policy. The government reimburses schools up to a cap of INR 17,640 per pupil per annum, an amount that might fall short of the actual fee for some schools. Intuitively, it is plausible that high-fee

1.7.2 Estimating the impact of attending elite private schools as a quota student

For the group of lottery winners, what is the impact of attending elite private schools as a quota student, relative to attending budget private schools? I estimate this using two-stage least squares framework on the sub-sample of lottery winners:

$$RTE\_Enrolled\_Elite_i = \alpha_1 WinningLotteryEliteAnyBin_i + X_i'\alpha_2 + \sum_{x=1}^{50} \gamma_x d_i(x) + \epsilon_i \tag{1.3}$$

$$Y_i = \beta_1 RTE\_\widehat{Enrolled}\_Elite_i + X_i'\beta_2 + \sum_{x=1}^{50} \delta_x d_i(x) + e_i \tag{1.4}$$

where, $RTE\_Enrolled\_Elite_i$ is the indicator that child i attends an Elite private school as a quota student, $WinningLotteryEliteAnyBin_i$ is the indicator that child i won the lottery at an Elite school in any bin, X_i is the vector of child and household characteristics, $d_i(x)$ are dummies taking a value of 1 if child i's estimated propensity score of winning a lottery at an elite private school lies in the respective 0.02 wide probability bin. As before, identification comes from within variation in lottery offers at elite or budget schools for groups of applicants who are otherwise similar in their ex-ante propensity of winning at elite schools.[43]

1.7.2.1 First stage

Table 1.7 shows the first stage - winning the lottery at an elite school, defined as schools lying above the 75^{th} percentile of the quality distribution, increases the likelihood of attending one, by 87pp - 88pp depending on the quality measure. The dependent mean shows the proportion of children enrolled at elite schools under the quota - this is 39% based on the

charging schools that charge in excess of what the government reimburses them, might increase their fee levels further, to compensate for the lost revenue per quota seat by extracting more revenues from the fee-paying students, keeping in mind the price elasticity of the fee-paying students. However, I cannot test this in the data. However, I do not think that schools increase their fee indiscriminately, in response to the policy - this is corroborated to some extent by the regression of school fees on the PCA index of school quality, which shows a strong positive correlation between the school quality index and school fees.

[43] The detailed step by step process of calculating these ex-ante propensities of winning at elite schools is explained in Appendix Section A.3.3.

Table 1.7: First stage of winning the RTE lottery at elite school on enrollment at an elite school

	RTE student at Elite school	
	Elite (PCA)	Elite (Fee)
	(1)	(2)
Won RTE lottery at Elite school	0.880***	0.869***
	(0.027)	(0.028)
Outcome mean	0.39	0.51
Control mean	0.00	0.00
Observations	1,019	973
R^2	0.85	0.82
Pscores of winning at elite	Yes	Yes
Controls	Yes	Yes
Avg quality (Elite=1)	4.37	43046.62
Avg quality (Elite=0)	2.34	12320.27

Notes: This table shows the first stage effects of winning the RTE private school lottery at an elite school on enrollment at an elite school as a quota student. The sample is restircted to lottery winners. Eliteness is defined using the 75^{th} percentile cutoff. Control variables include sex and age of child, indicators for father's and mother's education being greater than the respective means, indicators of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

PCA index measure, and 51% based on the fee-based measure. These differences stem from each measure identifying a different aspect of school quality and relatedly the fact that the same school might be categorized as elite based on one measure, but as budget on the other. It is also informative to learn about the average quality of schools that are categorized as elite versus those categorized as budget. Table 1.7 shows this for both the quality measures - the mean school fee for elite schools is about 3.5 times higher than that for budget schools, and this ratio is about 1.8 for the PCA index.[44]

1.7.2.2 Primary outcomes

Table 1.8 shows the LATE of attending an elite school on children's performance on phone-based assessments, using both measures of school quality. Treated compliers are children who attend elite private schools under the RTE quota because they won the lottery to an elite private school, and control compliers are those who do not attend elite schools under the quota because they lost the lottery to all the elite schools. Results show that elite schools increase English test scores on both the quality measures, however, there are no statistically

[44] Appendix Table A.15 shows how results vary when the percentile cutoff is changed to the 50^{th} and 90^{th} percentile.

Table 1.8: LATE of attending elite schools on performance on tests

	English	Math	English	Math
	Elite (PCA)		Elite (Fee)	
	(1)	(2)	(3)	(4)
RTE student at Elite school	0.699***	0.370	0.485**	0.138
	(0.270)	(0.267)	(0.242)	(0.240)
First stage F-stat	590.47	590.47	389.43	389.43
Outcome mean	0.06	0.04	0.06	0.05
Control mean	0.03	-0.05	-0.18	-0.20
Observations	318	318	303	303
R^2	0.14	0.13	0.20	0.17
Pscores of winning at elite	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on children's performance on phone-based assessments. The sample is restricted to lottery winners. As before, the number of observations is smaller here because the phone-based assessment on English and Math is available only for a subsample of lottery winners. Control variables include sex and age of child, indicators for father's and mother's education being greater than the respective means, indicators of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

significant gains in Math. As before I explore several mechanisms below.

1.7.2.3 Mechanisms

My survey data allows me to explore several school inputs such as, school's instructional modality, subjects taught and other school characteristics, and children's time use as potential mechanisms. I discuss these in depth in the subsequent paragraphs.

School inputs. While elite schools are no more likely to provide instruction in the two academic years (as measured on the extensive margin), they are however more likely to provide better instruction modalities. Table 1.9 shows that treated compliers are more likely to report receiving synchronous online classes (between 10 and 18 percentage points, based on the PCA and the fee-based measure, respectively), and less likely to receive text-based instruction (by 17 pp, based on the fee-based measure) during the period of remote instruction.

Table A.12 shows the differences in characteristics of elite and budget schools, and helps in understanding key differences across these schools. Controlling for the village fixed effects, elite schools are consistently more likely to have internet, more digital boards per pupil, more likely to be English medium, have a higher proportion of teachers trained in computers, a higher proportion of teachers with Bachelor's in Education degrees, and a higher proportion of general caste category students. The magnitude of differences in Bachelor's in Education degree is substantive, at 20 percentage points, indicating that elite schools are much more likely to hire teachers who have specifically trained to pursue a school teaching career.[45] These patterns suggest that the relative effectiveness of elite schools in remote instruction is likely to be a function of teachers being more qualified and also being more adept at dealing with digital technologies. In addition, these results are also robust to changing the elite cutoff to the 50^{th} and the 90^{th} percentile of annual fee (Table A.18 in Appendix), which provides a consistent story that elite schools were doing better in terms of providing instruction during the period of remote instruction.

Another channel that might explain gains in English relates to differences in the quality of English instruction, and differences in instructional time spent across subjects, across elite and budget schools. Appendix Table A.14 shows that while elite school goers are no more likley to be taught the conventional subjects (Math, English, Marathi and Hindi), they are more likely to have other subjects in their curriculum, such as General Knowledge, Arts, Music and Dance. These other subjects are taught in English (as suggested in balance table A.12 which shows elite schools being more likely to be English medium,) and this in turn might indirectly increase children's exposure to English thereby improving their test scores.

Children's time use. Finally, studying the impacts on children's time use (Table 1.10), I find that elite schools provide more hours of instruction per week (2.1 - 3.1 hours/week) relative to budget private schools.

Taken together, this evidence suggests some of the plausible mechanisms which might be at play and might matter for children's performance. While the lack of data on instructional time by subject precludes me for testing the role of that channel in explaining the results,

[45] In the Indian context, Bachelors in Education is a degree program that is specifically designed for those who aspire to become school teachers. It is typically a two-year program that one pursues after a three/four year undergraduate degree program, in order to become a school teacher.

Table 1.9: LATE of attending elite schools on school instruction

	Synchronous classes (online)	Recordings shared (audio/video)	Text-based activity plans (WhatsApp/SMS)	Synchronous classes (online)	Recordings shared (audio/video)	Text-based activity plans (WhatsApp/SMS)
	Elite (PCA)			Elite (Fee)		
	(1)	(2)	(3)	(4)	(5)	(6)
RTE student at Elite school	0.102*	0.024	-0.064	0.180***	-0.036	-0.174**
	(0.060)	(0.054)	(0.077)	(0.051)	(0.048)	(0.069)
First stage F-stat	1,151.81	1,151.81	1,151.81	1,129.44	1,129.44	1,129.44
Outcome mean	0.83	0.13	0.53	0.83	0.13	0.53
Control mean	0.77	0.15	0.59	0.69	0.17	0.60
Observations	1,005	1,005	1,005	959	959	959
R^2	0.09	0.05	0.12	0.18	0.08	0.13
Pscores	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on school's instruction modality and children's time use in educational activities. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

the results on differences in school characteristics, in terms of their baseline digital facilities, teacher quality, overall time spent in school, and likelihood of studying specific subjects provides some understanding of why elite schools were more effective in providing remote instruction and also enhancing children's learning.

Table 1.10: LATE of attending elite schools on children's time use

	School (hrs/week)	Tuition (after school) (hrs/week)	Homework (hrs/day)	School (hrs/week)	Tuition (after school) (hrs/week)	Homework (hrs/day)
	Elite (PCA)			Elite (Fee)		
	(1)	(2)	(3)	(4)	(5)	(6)
RTE student at Elite school	3.185***	-1.936**	0.007	2.181**	-0.314	-0.140
	(1.176)	(0.957)	(0.113)	(1.066)	(0.869)	(0.103)
First stage F-stat	982.18	982.18	982.18	937.84	937.84	937.84
Outcome mean	13.58	4.43	1.51	13.57	4.49	1.52
Control mean	12.98	4.91	1.51	12.38	5.00	1.48
Observations	1,019	1,019	1,019	973	973	973
R^2	0.10	0.09	0.04	0.11	0.09	0.03
Pscores of winning at elite	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β_1^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on school's instruction modality and children's time use in educational activities. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

1.8 Robustness checks

I perform a number of robustness checks to validate my findings.

1.8.1 Excluding applicants who are age-eligible to re-apply for RTE

The eligibility to apply for grade 1 admissions under the RTE in 2020-21 was that one had to be born no earlier than July 2013, and no later than October 2014. This means that the eligibility criteria to apply for grade 1 applications in the following academic year of 2021-22 was that one had to be born between July 2014 and October 2015. Hence, the age-eligibility to apply under the RTE policy for grade 1 spans more than a year, which implies that the very young applicants who applied for grade 1 lotteries in 2020-21, would be age-eligible to apply again under the policy, in the following year (academic year 2021-22). Appendix Figure A.8 shows the distribution of the birth year-months in the population of the applicants.

Figure A.8 shows that my sample of surveyed applicants contains some children who are eligible to re-apply under the policy in the following academic year of 2021-22; these are the applicants who were born between July 2014 and October 2014. This leads to two concerns. The first concern is that among the young applicants (who are age-eligible to re-apply) those who lost the lottery, can wait to try again next year, however those who won the lottery in the first year, might accept and enroll. This is concerning because such applicants can cause selection bias and also affect the composition of the control group. Since the control group comprises non-quota children, the presence of such applicants might lead to the control group having some children who became quota beneficiaries in the following year.[46] The second worry is that these young applicants might be different from relatively older applicants (age-ineligible to re-apply) on unobservable dimensions, for example, parental motivation to apply or child ability, which might be simultaneously correlated with children's performance on tests, and might bias estimates. Thus, I address these issues by limiting my analysis to the subsample of those who are age-ineligible to re-apply in the following year. I find that the results on phone-assessments are robust to excluding young applicants. Results are shown in Table A.20.

1.8.2 Using school level values to measure outcomes

My results use survey data from household level reports on children's outcomes. However, some of these variables could be measured with error. For example, certain variables correspond to school-level information, like - does the child's school provide instruction; what is

[46] I define my treatment as enrollment as an RTE quota student where the quota receipt is based on the 2020-21 school year.

the modality of instruction at child's school; frequency of classes at child's school etc. The ideal scenario would be to obtain administrative data from schools on these school-level variables, however, since such data is not available/collected by schools, I rely on household-level reports for these school-related variables. It is possible that the household-level responses to school-level variables might have measurement error such that there are inconsistencies in how children attending the same exact school might respond to any given question about the school, which in turn might lead to biases in the estimates.[47] I attempt to address this potential noise in the household-level responses, by creating a new variable that captures the school-level unique responses to these questions. I do this by coding the value of the new variable as the response that was most frequently chosen by students attending the same school with the goal that these new variables have less noise.[48]

Results based on this cleaning of school level variables are shown in Appendix Table A.21. I see that the results are still robust - RTE schools are still statisically significantly more likely to provide school instruction in the academic year of 2021-22, they are still more likely to provide synchronous and live classes. The standard errors on the estimates have also shrunk which is a mechanical result due to a decrease in the variation in the new outcome variable. These robustness check strengthen the validity of the main results.

1.8.3 Varying the ex-ante propensity scores of winning

The identification in my estimates comes from using the within-variation in children who had a similar ex-ante likelihood of winning the RTE lottery but varied in their final lottery outcome. To do this I control for narrow bins of ex-ante propensity scores of winning by simulating the lottery algorithm. I show that my results are robust to increasing the number

[47]For instance, consider a scenario where a total of five children attend school A, and of these five children, four children respond by reporting yes to the question that asks whether school A was providing instruction in the previous academic year, while one child responds no to this question. This would be problematic in the regression where the outcome measures the binary indicator of school providing instruction.

[48]I clean this by making a new clean variable at the school level based on whether the proportion of children who answer yes to the question at a given school exceeds half. This is under the assumption that the reports of more than 50% of the students are less likely to be incorrect.

of narrow bins of these ex-ante propensities, or in other words reducing the bin width. Reducing the bin-width would lead to stricter within comparisons, comparing children who had a very similar ex-ante likelihood of winning under the lottery. Appendix Figure A.7 shows this in a coefficient plot which shows how the LATE coefficient on test scores changes as the bin-width is reduced.

1.9 Conclusion

This paper studies the impact of India's Right to Education Act, an affirmative action policy that targets children of school entry age, on children's educational outcomes. Given that India has one of the world's largest schooling systems, the scale of the policy is huge and impacts millions of disadvantaged children.

I leverage the lottery-based allocation of students to schools to estimate the causal impact of attending a private school as a beneficiary under this policy on children's educational outcomes. The context is that of remote learning during the period of pandemic-induced school closures, and I find that being a beneficiary under the policy insured disadvantaged children from the risk of non-enrollment, and thereby helped with grade progression and maintaining the right grade-for-age trajectory. I find that these gains in enrollment translated into gains in performance in English and improvements in children's English test scores by 0.18 SD units. Exploring mechanisms, I find that relative to the non-beneficiaries, beneficiaries are more likely to attend private schools of higher overall school quality, where the main language of instruction is English, that are more effective in providing synchronous remote education technologies, and have a longer school week.

Next, given that the private schools themselves are differentiated in quality, I focus on the beneficiaries to estimate the causal impact of attending elite or higher quality private schools, relative to budget, or lower quality private schools, by leveraging the randomization in offers to elite schools. As before, I find statistically significant gains in English as a result of attending elite schools, which suggests that a single estimate of private school premium might be misleading. As before, the mechanisms point to elite schools being better at providing remote instruction and having a longer school week. Baseline differences in school characteristics further show that elite schools have better digital technologies, have higher proportion of teachers with better qualifications, and higher proportions of teachers trained in computers, all of which may matter in making remote learing more effective. Overall, my results suggest that private schools, and especially those at the upper end of the quality distribution, are effective in adapting to, and providing remote learning and in doing so they also enhance children's learning.

While the policy is successful in delivering these gains, however, there are concerns about whether the applicant pool is representative of the eligible groups in the population. Given that the fallback option for the majority of lottery losers (about 65% of control compliers) is to enroll as fee-paying students in private schools, this points to the concerns of regressive selection among eligible groups. Romero and Singh (2023) focus on the aspect of mistargeting and regressive selection in RTE and find that various constraints (information, documentation, application complexity) prevent poor households from applying under the RTE policy, despite them having a high demand for private schools. Thus, this points to the neccessity of improving targeting under the policy such that the benefits can percolate to the ones who most need it.

Chapter 2

Sibling Spillover Effects of Affirmative Action Policies during COVID-19: Evidence from India

2.1 Abstract

This paper examines the sibling spillover effects of India's Right to Education Act (RTE) during the period of COVID-19 induced school closures, focusing on the educational outcomes of siblings of policy applicants. The RTE mandates all private schools to reserve 25% of the incoming seats in grade 1 for low socioeconomic status students. Using administrative data of households that applied under this policy and survey data on educational outcomes of applicant children and their siblings, I estimate the intent-to-treat effects (ITT) of being a sibling in a household where the applicant child won the grade 1 private school lottery. My findings indicate that younger siblings in winning households were less likely to be formally enrolled in school compared to their peers in losing households, but benefitted from increased access to remote learning resources provided by the older applicant child's private school. Additionally, there were no significant differences in parental monetary and time investments between siblings in winning and losing households. This study shows that well-implemented affirmative action policies during economic hardships, such as the COVID-19

pandemic, can act as a safety net not only for the targeted individuals, but also benefit non-targeted individuals and can mitigate long-term educational inequalities.

2.2 Introduction

Over the past several decades, India, along with many other low and medium income countries, has made substantial progress in achieving the target of universal basic education with over 95 percent of children in India between the ages of 6 and 14 years now in school (Pratham, 2019). However, the COVID-19 pandemic led to extended school closures thereby creating unprecedented challenges for children's education. Multiple studies have estimated the impact of the pandemic on the extent of the learning loss, and potential ways for recovery (Kilenthong et al., 2023; Guariso and Björkman Nyqvist, 2023; Chatterji and Li, 2021; Andrew and Salisbury, 2023; Alasino et al., 2024). In mitigating the adverse impacts of the pandemic on children's learning, government-led remedial education programs and affirmative action policies have been especially useful (Romero and Singh, 2022; Dessy et al., 2023; Singh, Romero and Muralidharan, 2024; Agarwal, 2024).

A prominent example in the Indian context is the Right to Education Act (RTE) policy, which mandates all private schools to reserve 25% of incoming grade 1 seats for low socioeconomic status students. Recent research shows that the RTE was helpful in insuring vulnerable children against the risk of non-enrollment, maintaining grade progression, and at the same time improving their learning outcomes (Romero and Singh, 2022; Agarwal, 2024). While we have an understanding of how the RTE affected the targeted individuals, we know little about whether and how it might impact the non-targeted individuals. For example, in credit-constrained households, if school inputs improve for one child in the household, then does it impact parental decisions for other children in the household? These questions are especially relevant in times of severe economic hardship, such as the COVID-19 pandemic, when disadvantaged households face disproportionately higher challenges in supporting their children's education.

In this paper, I study the sibling spillover effects of India's RTE policy, by examining the educational outcomes of siblings of the applicants under the policy. I combine the administrative data of applications to grade 1 private school admissions under the policy for the 2020-21 school-year in the state of Maharashtra, along with survey data conducted with a sample of applicant households, where I collect detailed data on children's educational outcomes for both the applicants of the policy and their siblings like enrollment, availability of remote learning inputs from school, and parental monetary and time investments. Using these data I estimate the intent-to-treat effects (ITT) of being a sibling in a household where the applicant child won the grade 1 private school lottery to examine whether winning such a lottery for one of the children in the household impacts other children's educational outcomes.

I contribute by providing one of the first estimates of the impact of this policy on the educational outcomes of siblings of applicants. My data correspond to the applications made for grade 1 private school admissions under the RTE 25% quotas for the 2020-21 school-year in the state of Maharashtra, and hence correspond to the period of COVID-19 induced school closures. I conducted phone surveys with a sample of applicant households during the middle of the following academic year to collect data on children's educational outcomes like their enrollment status, availability of remote learning inputs from school, and parental monetary and time investments in siblings. These data allow me to estimate the intent to treat effects (ITT) of being a sibling in a household where the applicant child won the grade 1 private school lottery to examine whether winning such a lottery for one of the children impacts parental decision-making about other children's educational outcomes.

My first finding focuses on siblings' enrollment status. Overall, being a sibling in a winning household yields statistically and economically null impacts on siblings' enrollment outcomes. However, since educational outcomes of younger siblings who are yet to enter formal schooling are more likely to be affected, I estimate the ITT effects by introducing heterogeneity by siblings' age groups. I find that in households where the applicant child won the grade 1 private school lottery under the RTE 25% quotas, young siblings of pre-primary schooling age were 16.2 percentage points less likely (pval=0.000) to be enrolled in school compared to their peers in losing households. This drop in enrollment, however, does not persist for long, is only observed for the 2020-21 school-year which was also the first year of the COVID-19 pandemic, and exists only for siblings aged 4 years old at the time.[1]

However, since this was a period of school closures with schools struggling to provide virtual alternatives to in-person teaching, looking at enrollment patterns in isolation may not necessarily be the best indicator for whether students were learning during this time. Due to the disruptions in schooling, merely being enrolled in school was not enough and access to educational resources, either through school or within the home, mattered more for children's learning and development. In order to understand this, I estimate the ITT effects of being a sibling in a winning household on siblings' likelihood of receiving access to educational resources from schools. This could be by virtue of either their own school providing remote instruction (if they are enrolled) or by virtue of the enrollment status of the older child (grade 1 RTE applicant in the household) who, if enrolled, might be receiving remote schooling inputs from their school.

My second finding relates to the above point and I examine whether siblings' access to educational resources varied by the winning status of the grade 1 RTE applicant in the household. I conduct two exercises to explore this. First, I examine the extent to which enrollment of 4 year old siblings translates into the likelihood of getting any instruction from their *own* school. I find that in the 2020-21 school-year, 4 year old siblings in winning households were 9 percentage points less likely to be without any school instruction from their own school – this is lower than the enrollment gap of 16.2 percentage points (in favor of siblings in losing households). This suggests that for a substantial majority of the siblings

[1]Another finding worth mentioning is the impact on enrollment for primary school entry-age siblings (aged between 6-8 years) in the 2021-22 school year, i.e., one year after the applicant children had applied under the RTE 25% quotas. While statistically indistinguishable from zero at conventional levels, the results on the impact on enrollment for primary-school entry age siblings suggest that winning households were more proactive in seeking admissions for younger siblings nearing primary school entry-age by 8.5 percentage points (pval=0.13).

in this age group who were enrolled, schools were struggling to provide any remote learning during the 2020-21 school-year. Second, I examine that regardless of the enrollment status of 4 year old siblings, does being in a winning household make them any more likely to get access to *any* schooling inputs during the 2020-21 school year – either from their own school (if they are enrolled) or from the older child's school (grade 1 RTE applicant in the household, if they are enrolled). I find that for 4 year old siblings in the 2020-21 school year, being in a winning household made them 11.2 percentage points more likely to access remote schooling inputs from either their own school or the RTE applicant's grade 1 private school (pval = 0.018).

Thus, even though parents in winning households were less likely to formally enroll younger siblings in school, they benefited from positive spillovers of being in a winning household as it increased the likelihood of having access to remote instruction that was arguably relevant for their age. The second exercise described in the previous paragraph serves two purposes. In addition to providing evidence on whether having a private school lottery winner in the household leads to a supply of educational resources for young children in the household, it also helps in understanding the potential reason behind why young siblings in winning households see a drop in their enrollment relative to their peers in losing households. Since parents typically make enrollment decisions for their children around the same time of the academic year, one possibility is that parents of grade 1 lottery winners might expect to get remote schooling inputs from the winning child's private school and thus may forgo school enrollment for their younger children, with the expectation that the older child's remote schooling inputs might come in handy in teaching the younger children in the household.

My third finding relates to educational inputs that children might receive from within their home, and I focus on two types of home inputs: parental time and monetary investments in siblings. In recent years, government schools have seen a sharp decline in enrollment rates and fee-charging private schools are increasingly becoming the default choice for Indian households (Kingdon, 2020). In such a scenario, one might imagine that free-of-cost private school education for one child in the household might ease the household budget in families of lottery winners and in turn lead to potential income effects on household's educational spending. However, if such an income effect exists, does parental spending vary across siblings in winning and losing households?

Agarwal (2024) finds that on the extensive margin, parents are about 6 percentage points more likely to increase monetary investments in winning applicants, but does not find any evidence on the intensive margin of spending. I examine the extensive margin of household expenditure on various educational categories for siblings (school fee, after school private tutoring, and other expenses on books and stationary) but do not find any evidence of disparity across winning or losing households. Even though my findings point to a substantial difference in enrollment rates for 4 year olds across winning and losing households, one reason why this may not be reflected in differences in school-fee related expenses is because government run free-of-cost *anganwadis* are likely to be a more popular option among disadvantaged households for their pre-primary aged children.[2] I also examine the amount of time that siblings receive from household members on educational lessons, but do not observe any disparity for young siblings across winning or losing households, and observe a slight increase in time investments (on the extensive margin) for older siblings in winning households compared to their peers in losing households.

My final finding relates to treatment effect heterogeneity in enrollment by baseline observable characteristics of households. Given that the COVID-19 pandemic slowed economic activity and caused widespread health risks and unemployment, it is possible that other household characteristics might also play a role in how parents made educational decisions for their children. Using data on applicant households' RTE eligibility criteria (low income or disadvantaged caste), I find that in the 2020-21 school year, also the first year of the pandemic, enrollment rates for 4 year olds in winning households dropped for both low income

[2] Anganwadis are child care centers focusing on child nutrition, health check-ups, immunization and also serve as preschools, offering early childhood education to young children.

and disadvantaged caste households. However, compared to the disadvantaged castes, the drop was much more prominent for lower income households and also persisted in the following school year, suggesting that lower income families took longer to cope with the pandemic. Finally, while India has gone a long way in eradicating gender inequality in school enrollment, I find a strong evidence pointing towards gender-inequality in enrollment for young siblings during the time of the pandemic. My findings show that among 4 year old siblings, boys were 52 percentage points more likely to be enrolled in school relative to girls, across winning and losing households. Even though this effect only exists during 2020-21 school year (the first year of the pandemic), and fades out in the following school year, it goes on to suggest that young girls in disadvantaged families were disproportionately affected as a result of the pandemic.

My contributions to the literature are threefold. First, I contribute to the literature on parental decision-making about educational investments in children. Almond and Mazumder (2013) provide a review of papers examining how investments depend on children's endowments and whether parents reinforce or compensate based on these endowments. The findings provide mixed evidence with some studies indicating reinforcement (Adhvaryu and Nyshadham, 2016) and others suggesting compensatory investments (Bharadwaj, Eberhard and Neilson, 2018). Berry, Dizon-Ross and Jagnani (2020) provide experimental evidence suggesting that parents display strong preferences for equality in investments across children. I add to this literature by studying how endowments of a given child in the household – as measured by a positive schooling shock for the older child – might influence parental investments in other children in the household.

Second, I contribute to the literature on the impact of the COVID-19 pandemic on children's education and potential ways to mitigate these losses. Recent research from both lower-income and higher-income countries provide evidence on the adverse impact of the pandemic on children's education, and recommends post emergency programs (Ardington, Wills and Kotze, 2021; Azevedo et al., 2021; Beam et al., 2021; Engzell, Frey and Verhagen, 2021; Cattan et al., 2021; Maldonado and De Witte, 2022; Wolf et al., 2022; Guariso and Björkman Nyqvist, 2023; Kilenthong et al., 2023). I contribute by providing evidence of how well-implemented affirmative action policies can act as a safety net for the disadvantaged during times of severe disruption.

Finally, I contribute to the literature on spillover effects in education. Opper (2019) shows how peer-interactions between students leads teacher value-added to extend beyond the non-targeted students. A growing literature provides evidence that older siblings have a strong influence on younger siblings' educational choices (Joensen and Nielsen, 2018; Nicoletti and Rabe, 2019; Aguirre and Matta, 2021; Altmejd et al., 2021). While these papers focus on how older siblings influence major choice decisions made by younger siblings, I extend the literature by studying how parental decisions about younger siblings' education are impacted by the older child. Additionally, I examine how policies in general might have unintended effects on non-targeted individuals.

2.3 Background and Policy

The Right to Education (RTE) Act was enacted by the Indian government in 2009, and made education a fundamental right of every child aged 6-14 years. I focus on Clause 12(1)(c) of this act under which all private schools in India are mandated to reserve at least 25% of the seats in entry-level grades for children belonging to low socioeconomic (SES) families.[3] Children who get admitted to private schools under this policy are eligible to get free education from the respective schools until they complete grade 8. The government reimburses private schools to cover the school's tuition fee for children admitted under the quota.

I study the impact of this policy on the educational outcomes of applicants' siblings in the state of Maharashtra, India. The state adopted this policy in 2010 and the eligibility criteria

[3] Religious and linguistic minority schools are exempted under the RTE Act. Entry level grades comprise grade 1 and pre-primary grades (for example, nursery or kindergarten). However, in my context the bulk of applications comprise those made for admission in grade 1.

includes children from historically disadvantaged caste groups, low income backgrounds, and children with disabilities.[4] It adopted a centralized online application system under this policy, in the academic year of 2017-2018. The online application to apply to schools under this policy begins in the month of February and is open for a month, following which the allocation of students to schools begins based on a centralized lottery algorithm. The majority of schools in the state follow the June to April school year (some follow the May to March school year). Agarwal (2024) explains the policy and the lottery algorithm in greater detail.

2.3.1 Schooling in India

In India, formal schooling comprises early childhood education (ages 3-5 years), followed by lower primary education covering grades 1-5 (ages 6-10 years), upper primary grades 6-8 (ages 11-13 years), lower secondary grades 9-10 (ages 14-15), and finally upper secondary grades 11-12 (ages 16-17). While this reflects the average pattern for the right grade-for-age, there is no law that requires schools to follow this, and in practice, schools vary in their adherence to these norms. Early childhood education providers range from government-run anganwadi centers, stand-alone private preschools, preschools integrated with existing private schools. Parents choose between government and private schools for children's education starting from the primary grades.

As part of the school curriculum of early childhood education (ages 3-5 years), children are introduced to foundational concepts in numeracy and literacy. Foundational numeracy involves identifying shapes and numbers, writing numbers, counting objects, comparing less versus more. Foundational literacy involves writing and identifying letters, objects related to each letter, phonetics, rhyming words, vowels, consonants etc. The same is also true for anganwadis which offer early childhood care and education.

[4] Historically disadvantaged castes include Scheduled Castes, Scheduled Tribes, and Other Backward Classes (OBC). Low income families are defined as those earning less than INR 100,000 per annum ($4746 in PPP).

During the COVID-19 pandemic, schools and pre-schools in India were recommended by the respective state governments to transition to various forms of remote instruction (both asynchronous and synchronous).[5] Similarly, anganwadis were also instructed to offer early childhood education to young children through online technology (Ministry of Women and Child Development, 2020). Thus, even though schools and pre-schools were closed for in-person instruction, they continued to provided remote instruction.

2.4 Data

I use a combination of administrative and survey data to measure spillover effects on educational outcomes of siblings. The administrative data provides me with details of the universe of children who applied to grade 1 private school lotteries under the Right to Education Act's 25% reservation policy, in the state of Maharashtra for the 2020-21 school-year. After one and a half years of applying for grade 1 admissions under RTE, I conducted phone surveys with a sample of applicant households to collect data on educational outcomes of the applicant child and the closest-in-age sibling of the applicant child. To be eligible for being part of the sibling questionnaire, siblings had to be within the age range of 4-18 years at the time of the survey (i.e. in Nov, 2021).[6] My sample size comprises a total of 1358 households for whom I have data on the educational outcomes of siblings and contains information on school enrollment in the two academic years (2020-21 and 2021-22), school type (private or public), whether they receive any instruction from their school in the two academic years, and parental investments (time and monetary).

[5] For example, see this newspaper article which highlights that government of Maharashtra permitted schools and preschools to continue online education amidst the lockdown `https://www.huffpost.com/entry/open-defecation-india_b_7898834`.

[6] In the case that two siblings were equidistant in age with the applicant child, the survey instrument randomly picked one sibling.

2.5 Estimation

I estimate the ITT effects on siblings' educational outcomes as a result of being in a household where the applicant child won the RTE grade 1 private school lottery. Applications for private school lotteries by the applicant children were made for the 2020-21 academic year, and my goal is to examine whether the event of winning the lottery impacts educational outcomes of siblings within the household in the 2020-21 and the subsequent, i.e. 2021-22 academic year. Since the event of the applicant child winning the private school lottery is random only after conditioning on the private schools chosen at the time of application, I control for the simulated ex-ante probabilities of the applicant child winning the lottery, which are a function of the endogenous school preferences that parents listed at the time of applying under the RTE. Essentially, this strategy allows me to make comparable comparisons in the enrollment likelihood of siblings who are in ex-ante similar households (who have similar endogenous preferences for schools, and live in the same neighborhood) but faced a randomization in whether the applicant child won or lost the grade 1 private school lottery. Hence it allows me to study the causal impact on siblings' educational outcomes as a result of being in a household where the applicant child won the private school lottery.[7]

I start by estimating the following equation:

$$Y_i = \alpha_1 + \alpha_2 WinningHH_i + X'\alpha_3 + \sum_{x=1}^{100} \gamma_x d_i(x) + \epsilon_i \tag{2.1}$$

where, Y_i captures the educational outcomes of sibling i (enrollment status, whether school provides instruction, parental time and monetary investments), $WinningHH_i$ indicates whether the applicant child in sibling i's household won the grade 1 private school lottery under RTE 25% quotas, $d_i(x)$ are dummies taking a value of 1 if the applicant child's (in sibling i's household) estimated propensity score of winning the private school lottery lies in the respective 0.01 wide probability bin, X_i is the vector of sibling characteristics like

[7]See Agarwal (2024) for a detailed discussion of the lottery algorithm under Maharashtra's RTE 25% quotas and calculation of simulated ex-ante propensity scores of winning.

sex and age, and household characteristics like indicators of father's and mother's education being greater than the mean, dummy of low income quota versus disadvantaged caste quota, SES index, dummies of caste categories, and religion. These covariates are added only to increase the precision of my estimates and the results are robust to excluding them.

Since children's age is likely to play a critical role in parental decision-making about their education, my preferred estimation is given by equation (2.2), as it gets at the causal impacts for siblings by specific age groups.[8] The age groups in equation (2.2) reflect siblings' age at the time of survey (i.e., Nov, 2021).

$$\begin{aligned} Y_i = \alpha_1 &+ \alpha_2 WinningHH_i + \alpha_3 Age[4] + \alpha_4 Age[5] + \alpha_5 Age[6, applicantage) \\ &+ \alpha_6 Age[11,12] + \alpha_7 Age[13,18] + \alpha_8 WinningHH_i * Age[4] \\ &+ \alpha_9 WinningHH_i * Age[5] + \alpha_{10} WinningHH_i * Age[6, applicantage) \\ &+ \alpha_{11} WinningHH_i * Age[11,12] + \alpha_{12} WinningHH_i * Age[13,18] \\ &+ X'\alpha_{13} + \sum_{x=1}^{100} \gamma_x d_i(x) + \epsilon_i \end{aligned} \tag{2.2}$$

2.6 Results

In this section I discuss the causal impact of being a sibling in a household where the applicant child wins the RTE private school lottery, on a range of educational outcomes of siblings.

2.6.1 Enrollment

Table B.1 shows the overall ITT effect of being a sibling in a winning household, on siblings' enrollment in the two academic years. I also show enrollment by whether it is any school or private school. The main results indicate no overall impact on enrollment of siblings across

[8] The reference age category in equation (2.2) includes siblings who are older than the applicant child and up to 10 years old. At the time of survey (in 2021), the mean, min and max age of applicants is 7.62, 7.05 and 8.34 years, respectively.

winning or losing households across the two academic years. However, these estimates are weighted averages of the effects across siblings in various age groups. Since enrollment decisions are more likely to get impacted for younger siblings who are yet to formally enter schools, and are nearing pre-primary or primary grade entry-age, I look at treatment effect heterogeneity with respect to siblings' age groups.

Table B.2 shows the results taking into account heterogeneity by age, and I find that siblings who were 5 years old in 2021 were 16.2 percentage points less likely to be enrolled in school in the previous academic year when they were 4 years old, in winning households relative to losing households. This effect is highly statistically significant, and only present during the 2020-21 school year, which was also the first year of the COVID-19 pandemic, and goes away in the following school-year. This suggests that enrollment decisions of parents for their pre-primary aged children were impacted by whether their older child won or lost the grade 1 private school lottery. A potential reason behind such a sharp drop in enrollment for 4 year olds in winning households could be because parents of winners might forgo enrollment for their pre-primary aged children in expectation that the older child's private school would provide online learning materials which the household could use to teach their younger children.

While a drop in enrollment for young siblings of pre-primary school age suggests negative spillover effects as a result of being in winning households, however, further analysis is important to be able to comment on whether the spillover effects were truly negative. In the next section I explore this by examining the extent to which enrollment translates into access of educational resources from school.

2.6.2 Access to educational resources

Given that the context of this paper is school closures at the time of the COVID-19 pandemic, I argue that focusing on enrollment rates in isolation is likely to depict an incomplete picture about student learning and development. Due to school closures, schools had to adapt to provide virtual alternatives to in-person teaching. However, many schools struggled to provide remote learning during this time (World Bank, 2021; Agarwal, 2024). This meant that access to educational resources and children's engagement with the same was a much more important factor in determining whether students were learning. In my survey data, for each child who was a part of the survey (applicant and sibling), I asked parents about whether the child's school provided any instruction in the two academic years (if they were enrolled). These data help me in understanding two things. First, it helps in understanding the extent to which a sibling's enrollment translated into getting access to remote schooling inputs from their own school. Second, since instruction during this time was remote and hence likely to be available as a household public good, I examine whether being in a winning household (of grade 1 private school lottery winners) compensated younger siblings for the lack of their own school inputs due to non-enrollment.

Table B.4 presents these results for the academic year 2020-21. The first column looks at the extent to which sibling's enrollment translated into getting access to any instruction from their own school. The outcome takes a value of one for siblings who are enrolled and report getting instruction from their school, and takes value zero for those who are either not enrolled or enrolled but fail to receive any instruction from their school. The coefficients of interest for siblings by age groups are presented in Table B.5, and show that for 4 year old siblings in 2020-21, the likelihood of receiving school instruction in losing households goes up by 9 percentage points. However, this is substantially lower than the gain in enrollment likelihood (16.2 percentage points) experienced by the siblings of losers in this age group, compared to those in winning households. This implies that being enrolled did not necessarily mean that one was receiving school inputs during this school year.

The second column in Table B.4 looks at sibling's access to educational resources in the household, regardless of whether it comes from their own school or from the older applicant's school (grade 1 RTE applicant in the household).[9] As before, Table B.5 shows the coefficients

[9]This outcome variable takes a value of one if either the sibling own's school provides instruction or if the RTE applicant child's school provides instruction, and takes a value of zero otherwise.

of interest for siblings by age groups. I find that being in a winning household indeed compensates young siblings aged 4 years old in 2020-21, by increasing their access to remote schooling inputs by 11 percentage points, relative to the losing households. Taken together, these results show that being in a winning household plausibly led to positive spillovers for young siblings by facilitating access to virtual learning alternatives from the school of older child (grade 1 RTE applicant in the household). Additionally, Agarwal (2024) shows that the beneficiaries of the RTE 25% quotas were more likely to get synchronous online instruction, received longer hours of instruction per week, and had lessons on a wider variety of subjects, suggesting that younger siblings in the household might have had the opportunity to engage and take advantage of these virtual schooling sessions which were relevant for their age.

In the next sections I further study whether these effects vary by household socioeconomic status and parental education and find interesting heterogeneity on these dimensions.

2.6.3 Parental investments

Next, I examine whether the event of winning or losing the RTE lottery for grade 1 children led parents to invest differently in other children in the household. Since winning the lottery to a private school might free-up household resources and relax household budget constraints, households might either re-allocate these extra resources towards educational spending, or might devote them to food and household consumption.

I examine the impact on the extensive margin of monetary investments in siblings' education (school fee, private tutoring fee, other expenses), and don't find any evidence of differential investments in siblings across winning and losing households (Tables B.6 and B.7). While there is an increase in enrollment for young siblings in losing households, it is not that enrollment has increased in private schools, which in turn explains the lack of an effect when the outcome of interest is the binary indicator of expenditure on sibling's school fee. This makes sense because government run *anganwadis* are the biggest caterer of pre-primary education to disadvantaged families in India. Other providers like stand-alone private pre-schools and pre-schools integrated with existing primary schools are mostly used by the relatively affluent households as they typically charge high fees.

I also look at time investments in siblings which captures the extensive and intensive margin of the time that household members might be spending in helping them with educational activities or homework (Tables B.8 and B.9). While I do not see any evidence of differential time investments in young siblings across winning and losing households, I find that older siblings (13-18 years) in losing households are more likely to receive help, relative to their peers in winning households.

2.6.4 Treatment effect heterogeneity by age and baseline observables

In this section I explore whether the likelihood of siblings' enrollment varies by baseline characteristics of children and household. I focus on two dimensions: household's eligibility criteria to apply under RTE grade 1 admissions (indicator of whether the household applied under the low income category or the disadvantaged caste category) and sibling's gender. My goal is to examine whether the causal impact of being a sibling in a winning household is different for low income households, relative to disadvantaged caste households; and whether it differs based on the sibling's gender. I estimate these by modifying equations (2.1) and (2.2), and incorporating additional interaction terms.[10]

2.6.4.1 By eligibility criteria: low income versus disadvantaged caste

The pandemic brought about major economic inactivity and job losses. The extent of public health disaster was also disproportionately borne by the relatively non-affluent sections of the society. Thus, it is likely that household's socioeconomic status might also impact parental decision-making regarding children's educational outcomes. I examine whether the impact

[10] While I present the overall effects which do not incorporate heterogeneity by age, my preferred estimation is given by the following equation which incorporates both the heterogeneity by baseline observable

of being a sibling in a winning household varies based on whether the household is a low income household or a disadvantaged caste category.[11]

Tables B.10 and B.11 show these results without and with age heterogeneity, respectively. Table B.12 shows the coefficients of interest, by sibling age categories. I find that in the first year of the pandemic (2020-21), the drop in enrollment for siblings (who were 4 years old in 2020) in winning households was much higher in magnitude for low income households (drop of 21 percentage points) relative to the drop in enrollment for households belonging to the disadvantaged caste category (drop of 12.8 percentage points). Furthermore, specifically for low income households, the negative enrollment effect continued to persist for siblings in this age group (aged 4 years old in 2021) even in the second year of the pandemic (2021-22), and they were still less likely to be enrolled by 25 percentage points.

2.6.4.2 By sibling's gender

In a developing country like India, children's gender might play an important role in parental decision-making regarding their children's education. This is especially true in the context of a pandemic, when disadvantaged families might be more credit constrained than usual. Tables B.13 and B.14 look at whether the impact on enrollment as result of being in a winning household varies by the gender of the sibling. Table B.15 shows the coefficients of interest, by age. I find substantial disparities by child's gender.

characteristics (denoted by binary variable Z_i) and sibling's age:

$$\begin{aligned} Y_i = \beta_1 &+ \beta_2 WinningHH_i + \beta_3 Z_i + \beta_4 Age[4] + \beta_5 Age[5] + \beta_6 Age[6, applicantage) + \beta_7 Age[11, 12] \\ &+ \beta_8 Age[13, 18] + \beta_9 WinningHH_i * Z_i + \beta_{10} WinningHH_i * Age[4] \\ &+ \beta_{11} WinningHH_i * Age[5] + \beta_{12} WinningHH_i * Age[6, applicantage) \\ &+ \beta_{13} WinningHH_i * Age[11, 12] + \beta_{14} WinningHH_i * Age[13, 18] + \beta_{15} Z_i * Age[4] \\ &+ \beta_{16} Z_i * Age[5] + \beta_{17} Z_i * Age[6, applicantage) + \beta_{18} Z_i * Age[11, 12] + \beta_{19} Z_i * Age[13, 18] \\ &+ \beta_{20} WinningHH_i * Age[4] * Z_i + \beta_{21} WinningHH_i * Age[5] * Z_i \\ &+ \beta_{22} WinningHH_i * Age[6, applicantage) * Z_i + \beta_{23} WinningHH_i * Age[11, 12] * Z_i \\ &+ \beta_{24} WinningHH_i * Age[13, 18] * Z_i + X'\beta_{25} + \sum_{x=1}^{100} \gamma_x d_i(x) + \epsilon_i \end{aligned} \tag{2.3}$$

[11] Certain households qualify to be eligible under both the low income and caste category, however, they must choose only one category during the RTE application. Hence, I do not observe this detail in the data.

In the first year of the pandemic, regardless of the sibling's gender, I find that young siblings who are 4 years old at the time were less likely to be enrolled in winning households. However, for siblings nearing primary grade entry age (ages 5 and up in 2020), there is substantial disparity in school enrollment across girls and boys. In winning households, female siblings in this age group were much less likely to be enrolled in school (by 34.6 percentage points) but male siblings were more likely to be enrolled (by 19.6 percentage points). These results suggest that sibling's gender did play a role in parental decision-making about their education, and indicate that girls were at a disadvantage relative to boys. These effects however do not persist, and I don't observe any heterogeneity by sibling's gender in the following school-year.

2.7 Conclusion

This study provides one of the first estimates of the sibling spillover effects of India's Right to Education Act (RTE) during the COVID-19 pandemic, focusing on the educational outcomes of siblings of policy applicants.

My findings indicate that younger siblings in households where an older child won the grade 1 private school lottery were less likely to be formally enrolled in school during the 2020-21 school year. This trend appears to be a response to parents' expectations of benefiting from remote learning resources provided by the older child's private school. Despite this drop in formal enrollment, these younger siblings in winning households had increased access to potentially higher quality remote educational resources from the older child's private school, suggesting that the overall learning environment may still have been positively influenced.

There were no statistically significant differences in parental monetary and time investments between siblings in winning and losing households. This suggests that while the RTE policy alleviates some financial burdens, it does not necessarily alter the distribution of parental resources among children. Finally, the study highlights important heterogeneity in the impact of the RTE policy. Specifically, the negative enrollment effects were more pronounced among low-income households and for girls, highlighting the need for targeted interventions in mitigating potential long-term disparities.

Overall, my findings show how policies may have unintended spillover effects beyond the targeted individuals. Despite the observed decline in formal enrollment for younger siblings in winning households, the spillovers of the RTE 25% quotas were not necessarily negative. The policy mitigated potential adverse effects on their educational development, by facilitating positive educational spillovers within households.

Chapter 3

Women's Inheritance Rights and Household Sanitation

3.1 Abstract

Existing research indicates that females disproportionately benefit from having access to private toilets. However, lack of awareness regarding health-risks of improper sanitation, and limited intra-household decision-making power among females can be significant barriers to adopting toilets. In this paper, we study a novel link between household sanitation and policies that empower females. We estimate the causal impact of amendments to the Hindu Succession Act, which aimed to improve female property inheritance rights in India, on the presence of a toilet in their marital households using a difference-in-differences framework, and allowing for dynamic and heterogeneous treatment effects. Our findings show that the policy had a positive impact on the presence of a toilet in treated women's marital households, with the effects being concentrated in the states that adopted the policy late. The effects are primarily concentrated for women who were relatively young at the time of policy amendment. We attribute our results on increased toilet coverage to the policy's role in enhancing women's years of educational attainment and weakly increasing their intra-household bargaining power. Our paper highlights that policies empowering women's rights

can have unintended benefits on their socioeconomic status and can effectively improve sanitation coverage in regions grappling with open-defecation issues.

3.2 Introduction

Open defecation is a widespread problem in low and middle income countries and has been linked to illnesses and developmental problems like diarrhea and stunting in children, among many others. The practice is particularly prevalent in India, which accounts for 60% of the world's open defecation (Census 2011). The barriers to demand for toilets in India stem from deep-rooted cultural norms of religious purity, casteism, taboos surrounding menstruating women, as well as a widespread lack of awareness about the importance of sanitation. However, within the household, the presence of a toilet disproportionately benefits females (Aid Water, 2013; Jadhav, Weitzman and Smith-Greenaway, 2016). This is because females are often victims of sexual harassment when they go out in the open to defecate, urinate or attend to their menstrual hygiene (Jadhav, Weitzman and Smith-Greenaway, 2016; Hossain, Mahajan and Sekhri, 2022).[1] In spite of such difficulties faced by females, there exist several deterrents in the demand for toilets. First, given their low intra-household bargaining power, females are rarely the primary decision makers within their households (Coffey et al., 2014) and thus might be at a disadvantage to advocate for their needs. Second, lack of education and awareness about the importance of sanitation, is also an important factor in limiting the take-up of toilets (Coffey et al., 2014). This leads us to ask the question: do policies that are aimed at empowering women lead to an increase in the demand for toilets, a household public good that females value disproportionately more than males (Khanna and Das, 2016; Stopnitzky, 2017; Augsburg, Malde, Olorenshaw and Wahhaj, 2023)?

We study this question in the context of India's Hindu Succession Act of 1956 (HSA), which governed the property inheritance rights of Hindus, Sikhs, Jains, and Buddhists, constituting about 86% of the country's population. As with most personal laws, property

[1] Also see https://www.huffpost.com/entry/open-defecation-india_b_7898834

inheritance laws in India are also governed by religion, and the HSA established rules for the division of household property among heirs, in the event of unwilled succession. However, these rules were gender-unequal, granting substantially greater inheritance rights to sons than to daughters. In 1976, Kerala addressed this inequality, followed by Andhra Pradesh in 1986, Tamil Nadu in 1989, and Karnataka and Maharashtra in 1994. These states took measures to address the gender-inequality in HSA by equalizing the inheritance rights of daughters and sons before the national amendment in 2005, when all states eliminated the gender-inequality. However, for the five states that passed the amendment earlier, it did not benefit all females in these states; it only applied to and benefited those females who were unmarried at the time of the passing of the amendment, thus creating variation in treatment within the treated states.

We leverage this within-state variation across marital cohorts of females along with the staggered adoption of the HSA amendments (henceforth, HSAA) across states to estimate the causal impact of the HSAA in a difference-in-differences framework. Exploring the unintended benefits of female-empowering policies on their socioeconomic outcomes, our objective is to study whether the HSAA increased the likelihood of a toilet being present in the marital households of females. We follow the approach of Callaway and Sant'Anna (2021) and compare toilet ownership rates in marital households of women in treated states relative to untreated states, across women who got married after the passing of the amendment in their state (and thus benefited under the amendment), relative to those who were married at the time of passing of the amendment (and thus did not benefit under the amendment). Our identification assumption is that, in the absence of the HSAA, the likelihood of the presence of a toilet in treated states would evolve in parallel to states where the HSA was not amended, across marriage cohorts.

An obstacle in estimating the impact of the HSAA is that the treatment group is not perfectly observed in most publicly available datasets. While one of the eligibility criteria in order to benefit under the HSAA required females to have been unmarried at the time of passing of the amendment in her state, another condition was that her natal household property[2] must have been undivided at the time of the passing of the amendment in her state.[3] While the first condition is observed, the second condition is unobserved in nationally representative survey datasets since data on such natal household characteristics of married women in household are typically not asked in survey datasets.[4] Due to these data limitations, most studies in the literature have ignored this data caveat, with the exception of Roy (2015) and Deininger, Goyal and Nagarajan (2013), who use timing of death of grandfather and father, respectively, as a proxy for timing of household property division.[5] We address this common data caveat by formally showing that under generic assumptions, we can identify and estimate lower bounds of the true average treatment effect on the treated within a difference-in-differences framework, even while allowing for heterogeneous and dynamic treatment effects in a staggered policy adoption setting.

Allowing for heterogeneous treatment effects, we find that following the amendment to the HSA, women who were eligible under the amendment were on average at least 4-6 percentage points more likely to have a toilet in their marital household, relative to those who were not eligible. These effects are observed only for the late adopting states of Maharashtra and Karnataka that passed the amendment in 1994, and particularly for the cohorts of women who were young at the time of policy amendment. We find neither statistically significant nor economically meaningful impacts in the early adopting states. Exploiting variation in treatment across marriage cohorts over time, we explore the dynamic treatment effects of the

[2]In the context of India, "natal household property" refers to the property owned by a woman's family of birth, typically including assets such as land, which may be subject to inheritance laws.

[3]This is because the amendment did not apply retrospectively. If a household's property was already divided before the amendment was passed in the state, then the daughters of that household were not eligible to receive their notional share of the property even if they satisfied all other eligibility criteria.

[4]One reason is that marriages in India are patrilocal, meaning females move to their husband's natal household after marriage. Consequently, survey datasets mainly capture marital household characteristics, with limited data on the natal households of married women. The Rural Economic and Demographic Survey (REDS) is a partial exception, as it includes retrospective information on basic details about all household members, including married daughters.

[5]They use REDS which contains data on the timing of the grandfather's and father's death. However, REDS is not useful for our study since it lacks information on whether married daughters have a toilet in their marital households, our outcome of interest. To our knowledge, survey data on the timing of property division in India does not exist.

policy using an event-study design. Our findings reveal that the policy primarily impacted cohorts of females who were relatively young at the time of the amendment. This result is consistent with the existing literature on the homogenous treatment effects of the policy amendment on other outcomes of females (Roy, 2015). Our pre-period event study estimates along with pre-trend tests following Callaway and Sant'Anna (2021) provide no statistical evidence to suggest that the pre-treatment effects are statistically or economically different from zero, strengthening our identification assumption of conditional parallel trends.

Several papers have studied the impact of amendments to the Hindu Succession Act on various other female outcomes. Prior studies provide mixed evidence on HSAA's direct impact on improving women's inheritance rights,[6] but they consistently show that the policy led to alternative forms of parental investment, particularly in education (Deininger, Goyal and Nagarajan, 2013; Roy, 2015; Bose and Das, 2021; Ajefu et al., 2022).[7] Studies also show that the HSAA led to increased dowries (Roy, 2015), enhanced women's decision-making capacity (Deininger et al., 2019; Mookerjee, 2019; Biswas, Das and Sarkhel, 2024; Bose and Das, 2021; Ajefu et al., 2022), and increased labor market participation (Heath and Tan, 2014). Some studies examine the second-generation effects of HSAA and show improvements in child nutrition and health outcomes (Ajefu et al., 2022), but no impact on children's education levels (Bose and Das, 2021). For the purpose of our question, even if the amendment did not increase inheritance rights for daughters, alternate channels of investments such as increased years of education are likely to raise awareness of the importance of sanitation, which in turn is likely to promote behavioral change against these longstanding practices. Higher years of educational attainment are also likely to have increased women's intra-household bargaining

[6] Roy (2015) and Agarwal, Anthwal and Mahesh (2021) find that the amendments were not successful in improving actual inheritance received by women. Several studies suggest that the reason behind parental reluctance in bequeathing land (the main form of ancestral property in India) to daughters is due to patrilocality (the norm of daughters moving to their husband's house post-marriage) and the related risk that the property ends up being controlled by the in-laws of the daughters (Agarwal, 1994; Agarwal, Anthwal and Mahesh, 2021; Bhalotra, Brulé and Roy, 2020). Only Deininger, Goyal and Nagarajan (2013) find that the HSAA improved female inheritance.

[7] Unintended negative impacts of the policy have also been documented. For example, HSAA has been found to increase sex-selective abortion in areas with higher preference for sons (Rosenblum, 2015; Bhalotra, Brulé and Roy, 2020), increase suicide rates for both men and women (Anderson and Genicot, 2015).

power, a finding reported by some papers in the literature (e.g., Ajefu et al. (2022)).

To the best of our knowledge we are the first to examine whether female-empowering policies (such as the HSAA) increase the likelihood of household toilet ownership rates, which is our primary contribution.[8] Secondly, we address the typical data caveat in this literature of not observing an eligibility criterion that determines treatment status, and show that even without observing this in the data we can identify and estimate lower bounds on the true treatment effects parameters of interest under generic assumptions. Thirdly, we allow for heterogeneous and dynamic treatment effects in this setting with staggered policy adoption following the recent literature on treatment effect estimation using design-based difference-in-differences methods (De Chaisemartin and d'Haultfoeuille (2020) and Callaway and Sant'Anna (2021)).

We explore mechanisms that plausibly drive our main result, which is that the policy led to an increase in the likelihood of women's marital households having a toilet. Motivated by the literature, we investigate two main mechanisms: years of educational attainment of women and their intra-household bargaining power in their marital household, while allowing for heterogeneous and dynamic treatment effects of the policy. Our analysis shows that the HSAA boosted educational attainment for treated females in the late-adopting states, especially for those who were young at the time of the policy amendment. This increase in education likely raised awareness about the importance of sanitation in reducing health risks and challenged entrenched religious and cultural norms of open defecation, thus contributing to higher toilet adoption rates. While our findings on bargaining power are less conclusive, we observe some positive effects, suggesting weak evidence of the role of bargaining power in the take-up of toilets. Overall, our paper highlights that education is the primary driver behind the HSAA's success in enhancing take-up of toilets, and offers some evidence that

[8] In doing this, we also contribute to the literature that studies how identity and/or gender of the person receiving certain types of income might affect household outcomes. For example, Duflo (2003) finds increases in nutritional status of young girls when pensions are received by women, and finds no effect when pensions are received by men. The author suggests that efficiency of public transfer programs may depend on the gender of the recipient.

improved intra-household bargaining power might also play a role. This suggests that even for females, who have a disproportionately stronger preference for accessing private toilets compared to males, education might be crucial in overcoming the persistent and longstanding behavioral norms of practicing open defecation.

The rest of the paper is organized as follows: Section 2 describes the institutional background of the Hindu inheritance law in India. Section 3 outlines the data. Section 4 outlines the empirical strategy. Section 5 presents results followed by Section 6 which talks about mechanisms. Section 7 talks about robustness checks and Section 8 concludes.

3.3 Institutional Details

3.3.1 The Hindu Succession Act of 1956 (HSA)

Inheritance rights in India vary by religion. The Hindu Succession law of 1956 governed the property rights of Hindus, Sikhs, Buddhists and Jains. It established the rules of division of household property in the aftermath of the death of the patriarch of the family, in absence of a will.[9] Two major legal doctrines governing Hindu inheritance are the *Mitakshara* and *Dayabhaga* schools. The HSA governed the property rights following the *Mitakshara* system which distinguishes a person's individual property from joint ancestral property. Joint ancestral property includes that which was inherited patrilineally or any private property which was merged into the ancestral property or property acquired by the joint family, and primarily includes ancestral land (Agarwal, 1994; Rosenblum, 2015). Under HSA, only the male heirs (sons, grandsons, great-grandsons) were entitled to a share in this property. Separate property was accumulated and acquired separately and the owner had the freedom to bequeath it to whomever they wished. Under the original rules, daughters of a Hindu male dying intestate (i.e., without writing a will) were equal inheritors, along with sons, only of

[9]According to field studies, more than 65 percent of people in India die every year without making wills, and this proportion is much higher in rural areas, suggesting the importance and applicability of HSA in governing inheritances for individuals (Agarwal, 1994; Deininger, Goyal and Nagarajan, 2013).

their father's separate property but had no share in the joint property. Rights to the joint property were limited to the *coparceners*[10] that only constituted male members of the family. Since joint property typically takes the form of land that is generally family owned, females were at a significant disadvantage under the original inheritance rules and HSA was by no means a gender-neutral law.

3.3.2 State Amendments to Hindu Succession Act (HSAA)

Five states in southern India enacted legislation to amend the law at the state level, before the amendments were nationally ratified in 2005. Kerala in 1976, Andhra Pradesh in 1986, Tamil Nadu in 1989 followed by Karnataka and Maharashtra in 1994 took measures to redress the gender inequality inherent in the original HSA. Under these amendments, daughters were granted equal inheritance rights as sons in the joint property but this was conditional on daughters satisfying some eligibility criteria. The following conditions needed to be satisfied by daughters to be eligible under the HSAA. First, she had to reside in one of the five reform states at the time of the amendment. Second, she had to be unmarried at the time when the amendment was passed in her state. Third, she had to hail from one of the HSA religions (Hinduism, Jainism, Sikhism or Buddhism). Finally, the household property of the woman's parental household must have been undivided at the time of the passing of the amendment in her state. On September 9, 2005, all the eligibility criteria were removed, and the amendment was implemented at the national level granting equal claims to the joint household property to daughters and sons. The HSAA not only made daughters' status equal to that of sons' but, by the very definition of coparceners it also meant that her share in joint family property cannot be willed away by her father (Deininger, Goyal and Nagarajan, 2013).

[10] In the context of Indian inheritance laws, "coparceners" are family members who command equal shares in the inheritance of undivided ancestral property. Traditionally, coparceners included only male members of a family, before the amendments to the Hindu Succession Act were instated.

3.4 Data

We use data from the 2005-06 wave of National Family Health Survey (or NFHS-3), a large scale, cross-sectional and nationally representative survey of households across 29 states in India. It collects detailed information about the socioeconomic status of households, educational attainment for all household members, and an additional questionnaire for women aged 15-49 years. The questionnaire covers a variety of questions on the marital status of women, including year of marriage, as well as questions regarding women's autonomy and decision-making across various dimensions. The data also has information on toilet ownership in the marital household of women, which is our outcome variable of interest.[11] However, this is an eventual outcome observed only in the year 2005 (i.e., the year of survey), and we are unable to observe the year in which the toilet was constructed and whether it was before or after the women in our sample got married into these households. [12]

We define exposure to treatment in two alternate ways. First, we look at whether *any* married woman in a given household was exposed to HSAA. Since the majority of the sample consists of households headed by males, this includes the wife of the household head and any daughter-in-law(s) of the household. Second, we look at whether the wife of the household head was exposed to HSAA. Our goal is to estimate the impact of being exposed to gender-equal household property inheritance laws under HSAA on the presence of a toilet in the marital household of women. We do this because intra-household decision making power might vary with the presence of multiple women in the household.[13]

[11]NFHS has information on whether the household has access to a toilet facility, type of toilet facility (with or without flush, type of pit latrines, composting toilet etc.), and whether the household shares the toilet with other households. For our main analysis we code household toilet ownership as a "yes" if the household has access to any kind of toilet and code it as a "no" if they have no facility.

[12]Since we don't observe the year in which households get a toilet, we are unable to directly examine whether the household already had a toilet at the time the woman got married into the household, or after the woman married into the household. This matters in understanding the mechanisms at play. To get at this, we conduct an indirect test by analyzing an earlier round of the NFHS data. We discuss this issue in more detail in Section 3.7.

[13]Anukriti et al. (2020) and Anukriti et al. (2022) find the daughter-in-laws command a lower bargaining position within the household as compared to the mother-in-law, in the context of India. Calvi (2020) finds that women's bargaining position in Indian households weakens at post-reproductive ages. In the context of polygynous households in Sub-Saharan Africa, Hidrobo, Hoel and Wilson (2021) find that junior wives have

Following previous papers in the HSA literature, we restrict our analysis and sample to women belonging to one of the HSA-eligible religions—Hinduism, Sikhism, Jainism, and Buddhism– in order to restrict comparisons across treated and control groups within the eligible religions.[14] In another restriction, we drop the households belonging to the state of Kerala (one of the five states to pass the HSA amendment) because of two reasons. First, the amendment in Kerala abolished joint family property altogether, and the reform applied to all daughters regardless of their marital status (Agarwal, Anthwal and Mahesh, 2021).[15] This would imply that there is no within-state variation to identify the impact of the policy on any outcome in Kerala. Second, being one of the more progressive states, Kerala already has almost universal toilet coverage in the year prior to the amendment. This leaves us with a total of 28 states in our main analysis.

We restrict our time window such that the start and end years are 1980 and 2005, where years represent the year of marriage of women in the sample since that is the relevant time dimension. Given staggered adoption of HSAA across states (and removing Kerala for the reasons mentioned in the previous paragraph), Andhra Pradesh was the first state to pass the amendment in 1986, followed by Karnataka in 1989, followed by Tamil Nadu and Maharashtra which were the last two states to pass the amendment in 1994 before the national ratification in 2005.

In summary, our sample selection criteria are the following: since we are interested in whether women have a toilet in their marital household, we focus only on the married women in the household (which includes the wife of the household head, and any daughters-in-laws). We drop the state of Kerala, since the state has universal toilet ownership in the year prior to the amendment. We only keep women belonging to the HSA religions, and those who got

less decision-making power and receive fewer resources compared to senior wives.

[14]Another reason to not use non-HSA religions (Muslims, Christians, Parsis, and Jews) as a comparison group is because the non-HSA religions constitute about 14% of the data in nationally representative datasets and lead to power issues since our estimation strategy requires having data on women who got married in each year (between 1980 and 2005) across each state.

[15]Several papers in the HSA literature drop Kerala following this reasoning (Deininger, Goyal and Nagarajan, 2013; Rosenblum, 2015).

married after the year 1980.[16]

3.5 Empirical Strategy

We begin by discussing how—in our case with cross-sectional data—we are able to estimate the average treatment effect on the treated while allowing for heterogeneous treatment effects. At first glance, the limitation in implementing a difference-in-differences strategy in our setting arises from the lack of a panel, or even of repeated cross-section data. What enables us to allow for heterogeneous effects across groups, in spite of this seeming limitation, comes from the year of marriage component of the eligibility criteria, relative to the year of policy implementation across states.[17] This brings the dimension of time into our analysis and allows us to compare treated and untreated cohorts of women within a given state (as defined by whether they were unmarried or married by the year of policy implementation in their state). A potential challenge in estimating the unbiased effect of the HSAA is if households respond by strategically selecting into or out of the policy. Selecting into the policy might happen if families of daughters delay their marriage such that daughters are eligible for increased inheritance in anticipation of the policy. Selecting out of the policy, and thus excluding daughters from their rightful share in household property, might happen if parents marry off their daughters before state level amendments. Such self-selection will not lead to clean comparisons in the difference-in-differences framework and will produce biased estimates. We check for these selection patterns in the data by plotting the distribution of year of marriage and age at marriage and find evidence of no such patterns (Fig 3.5 and Fig 3.6).

Recent advances in the literature on treatment effects estimation in a staggered policy adoption design using two-way fixed effects have been documented to produce potentially misleading results when the treatment effects are heterogeneous across groups and/or over

[16]The latest year of marriage in our data is 2004.

[17]In our case, a group refers a given year of policy implementation. Hence each group comprises the set of states which pass the amendment in a given year.

time (Borusyak and Jaravel, 2018; De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021). Hence, we estimate the average treatment effect on the treated using methods proposed by Callaway and Sant'Anna (2021) which allow for heterogeneous treatment effects. For inference, we use wild bootstrap standard errors clustered at the state level allowing for arbitrary correlation between the unobservables within a state (our analysis contains 28 states or clusters).

Following Callaway and Sant'Anna (2021), we estimate the group-time average treatment effects of the policy on the treated. Let i denote a woman and let t denote the year of marriage of the woman (thus representing the cohort). Let G_i denote the group in which i belongs that represents the year of policy implementation in states where HSA was amended. G_i takes a value of zero for any i who belongs to the non-HSAA states (i.e., states that did not amend HSA before the national ratification of the Act in 2005), representing that these individuals were never treated.[18]

The outcome of interest is $Toilet_{igt}$ which equals 1 if woman i married in year t belonging to group $g \in \mathcal{G} \equiv \{1986, 1989, 1994\} \cup \{0\}$ has a toilet in her household at the time of the survey.[19] We report estimates using the never treated as the comparison group in our main analysis. Results are robust to using not-yet treated units as the comparison group instead.

3.5.1 Assumptions

We make the standard assumptions in Callaway and Sant'Anna (2021), namely, random sampling, sharp design, no treatment anticipation and conditional parallel trends in post-treatment periods based on the never-treated group. We rely on conditional parallel trends assumptions for the purpose of identification of the parameter of interest. This assumption (equation 3.1) essentially means that in absence of HSAA, the evolution of toilet ownership in the amendment states would be parallel to the toilet ownership in never-treated states,

[18] The notation in Callaway and Sant'Anna (2021) for never treated units i is $G_i = \infty$ denoting that these units are treated at time infinity.

[19] This is unlike standard outcomes in a difference-in-differences settings where the outcome is a realization at time period t. In our case the outcome is a point-in-time realization.

for households with similar characteristics (such that these characteristics are unaffected by treatment). To achieve this we control for the following household characteristics in our estimations which we think are relevant for household toilet ownership and are unlikely to be affected by the HSAA: wealth index, caste and indicator of urban residence.

A statement about the counterfactual, equation (3.1) says the following: the differences in average potential outcomes (toilet ownership in absence of policy) for any two cohorts of women that got married at any two years (t, t') in any amendment state would be the same as the difference in average outcomes for the same two marital cohorts in the non-amendment states.

$$\mathbb{E}\left[Y_{it}(0) - Y_{it'}(0) \mid X_i, G_i = g\right] = \mathbb{E}\left[Y_{it}(0) - Y_{it'}(0) \mid X_i, G_i = 0\right] \tag{3.1}$$

for all $t, t' \geqslant g_{\min} - 1$, where $g_{\min}$ is the first period where any married woman is treated (1986 in our case), and X_i denote time-invariant covariates of woman i. Equation 3.1 specifies that in absence of the policy for each group the potential outcomes between treated and never treated cohorts would evolve in parallel on average.

3.5.2 Average treatment effect on the treated

Under the assumptions mentioned in the previous section, variation in treatment timing relative to the year of marriage can be used to identify the average treatment effect on the treated for each group g (year of policy implementation) and time period (marriage cohort) t denoted by $ATT(g, t)$. Intuitively, we can identify $ATT(g, t)$ for each group g married in year t, by comparing the expected change in outcome between cohorts in a given group g that were married in year t and those that were married in year $g - 1$ (the year prior to policy amendment for the group) to the same difference for control states (never treated or not yet treated). Formally, under the conditional parallel trends assumption, using any comparison group $\mathcal{G}_{\text{comp}}$, the average treatment effect on the treated for each group g and time period t is given by:

$$ATT(g,t) = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid X_i, G_i = g\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid X_i, G_i \in \mathcal{G}_{\text{comp}}\right] \tag{3.2}$$

3.5.3 Data caveat and bounds on the true parameter

As discussed earlier, one of the eligibility criteria under the amendment was that the woman's natal household property should have been undivided at the time the amendment was passed in her state. However, our dataset does not have information on this condition. Roy (2015) studies the impact of the HSA amendment on educational attainment and dowries for daughters, but these are pre-marriage outcomes and the data used—the Rural Economic and Demographic Survey (REDS)—has retrospective information on all the members of the household including daughters who have married and left the household. REDS does not have post-marriage data for daughters who have left the household and hence it is not useful for the purpose of our analysis. The author uses data on the timing of a daughter's grandfather's death as indicative of whether the household property was undivided at the time of amendment in her state. This is because in Indian households, property typically gets divided when the patriarch of the family dies. However, we do not observe this as most publicly available surveys, including the one we use, do not ask this question for the married females residing in the household. Consequently, in our empirical model, the treatment group is likely to be mis-measured because of which some individuals who should ideally be in the control group might end up in the treatment group. This mis-measurement could lead to a bias in the estimated average treatment effect. We address this by deriving bounds on the true parameter when the treatment group is mis-measured.

For each group g we assume that the timing of division of property is independent of other variables. This is motivated by Roy (2015) who uses the death of the grandfather as the time which defines division of property which is plausibly a random event. This assumption ensures identification of lower bounds on the group-time treatment effects. Thus any aggregation of the group-time ATTs will result in a lower bound of the overall aggregated ATT. In particular, we can show that not observing one of the eligibility criteria defining treatment status can allow us to identify lower bounds of the treatment effect if the unobservable criterion is independent of other variables and can only affect the outcome through treatment. The intuition of this proof is as follows: When the researcher does not observe one eligibility criterion, but observes all other criteria—some individuals who truly belong to the control group (satisfying all other eligibility criteria except the unobserved one) end up in the treatment group instead. The treatment effect for these individuals should be zero. Therefore, by mistakenly increasing the number of treated individuals, the average treatment effect is reduced. Note that this also decreases the number of control group individuals. However since the true treatment effect for them should be zero, the average effect on the control group is unaffected. Hence, if the true treatment effect is positive, the estimated treatment effect will understate the true effect of the treatment, which is our case.

Formally, we can state this as follows:

Proposition 1. *Suppose for each unit i we only observe its group identity G_i, but we do not observe one criterion that determines treatment eligibility. Let us denote this unobserved treatment eligibility criterion as a dummy variable b_i which takes a value 1 if unit i is eligible for treatment. We continue to maintain standard assumptions of random sampling, no anticipation and parallel trends based on a comparison group $\mathcal{G}_{comp}$ (not-yet treated or never-treated) which identifies $ATT(g,t)$ for all groups $g \in \mathcal{G} \setminus \mathcal{G}_{comp}$ and all time periods t when all criteria of treatment eligibility are observed. Under an additional assumption that b_i affects potential outcomes of unit i through treatment only and is independent of other group identity, the $ATT(g,t)$ identified under this data limitation is a lower-bound on the true $ATT(g,t)$ for all groups $g \in \mathcal{G}$ and all time periods t. This also extends to the case where we condition on a set of covariates X_i which are independent of b_i and only affect potential outcomes through treatment.*

Proof. See Appendix Section C.3. □

We use the doubly robust estimator proposed in Callaway and Sant'Anna (2021) who extend estimators for two-period, two groups setup developed by Sant'Anna and Zhao (2020) to multiple periods and groups, to estimate the $ATT(g,t)$'s. The doubly robust estimator performs better than alternative estimands such as outcome regressions and inverse probability weighting, especially when the data are not a balanced panel, which is our case. See Callaway and Sant'Anna (2021) for more details.

3.6 Results

In this section we report and discuss the results from our estimation of the effect of the HSA amendments allowing for heterogeneous and dynamic treatment effects. As discussed above, we interpret our estimates as a lower bound of the true treatment effect. We show the results with both - any married woman, and wife of the household head. However, we prefer the specification that considers retrospective treatment exposure of any married woman in the household. This is because constructing a toilet is a household-level decision, and exposure to HSAA of any married woman in the household might play a role towards that household level outcome.

3.6.1 Heterogeneous Treatment Effects

We report the group-wise and the aggregated average treatment effects of the policy on the treated in Tables 3.1 and 3.2. We find evidence of heterogeneous treatment effects of the policy across the states that adopted the policy in different years. In particular, we find that the policy statistically significantly increases the likelihood of toilet ownership in the states that passed the amendment in 1994 by 4.7 to 6.3 percentage points. We find that the policy did not have a statistically significant impact on the likelihood of women's marital households having a toilet for the earliest adopting states in our sample—Andhra Pradesh, in 1986 and Tamil Nadu in 1989— with the corresponding estimates being very close to

Table 3.1: Impact of HSAA on likelihood of marital household having toilet (Any married woman)

	(1) Never treated	(2) Not yet treated
Aggregate ATT (GAverage)	0.0233	0.0243
	(0.0177)	(0.0177)
ATT of units treated in 1986	-0.00661	-0.00416
	(0.0352)	(0.0347)
ATT of units treated in 1989	-0.00151	0.000280
	(0.0340)	(0.0335)
ATT of units treated in 1994	0.0472*	0.0472*
	(0.0253)	(0.0253)
Observations	42,765	42,765

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). We present estimates by two different comparison groups: never-treated as comparison (column 1) and not-yet treated as comparison (column 2). Standard errors are computed using wild cluster bootstrap at the state level.

zero. These results hold true regardless of how we assign the treatment status of households. Tables 3.1 and 3.2 respectively show these results when the treatment is defined by whether any married woman and whether the wife of the household head in a given household was exposed to HSAA. To put the results into perspective, Geruso and Spears (2018) find that a reduction in open defecation by 10 percentage points is associated with a decrease in infant mortality by 6 per 1000 live births. We also conduct additional analysis that restricts the sample to rural households and find similar results (Appendix Table C.4).

3.6.2 Average treatment effects on the treated over time

We estimate an event-study framework to investigate the average treatment effects of the policy on the treated over time by comparing outcomes of marriage cohorts across states over time. This exercise is useful in shedding light on how the policy impacted different cohorts of women. In particular, for each group (defined by year of policy implementation) and time period (defined by year of marriage) the average treatment effect on the treated is estimated

Table 3.2: Impact of HSAA on likelihood of marital household having toilet (Wife of household head)

	(1) Never treated	(2) Not yet treated
Aggregate ATT (GAverage)	0.0357*	0.0375*
	(0.0198)	(0.0198)
ATT of units treated in 1986	0.00734	0.0113
	(0.0393)	(0.0387)
ATT of units treated in 1989	0.00690	0.0105
	(0.0373)	(0.0366)
ATT of units treated in 1994	0.0634**	0.0634**
	(0.0287)	(0.0287)
Observations	32,169	32,169

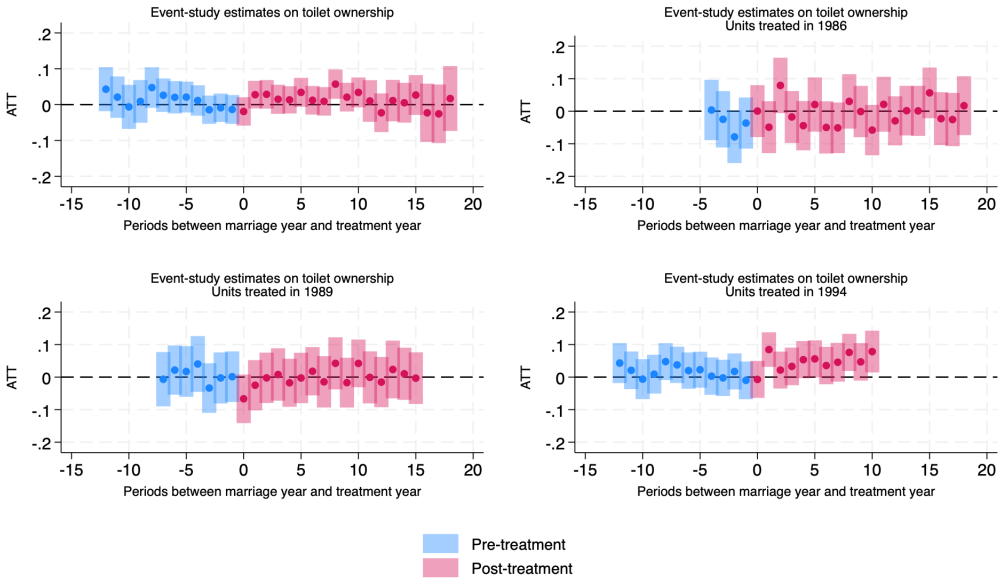
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). We present estimates by two different comparison groups: never-treated as comparison (column 1) and not-yet treated as comparison (column 2). Standard errors are computed using wild cluster bootstrap at the state level.

by comparing differences in average outcomes of the group in the given time period relative to its average outcome in the time period prior to policy implementation in that group, with that of the comparison group's differences in average outcomes for the same pair of time periods. The event study framework additionally provides estimates of the treatment effect of the policy for the cohorts that got married before the policy was implemented in their state, thus providing a test of the identification assumption of conditional parallel trends. We plot the event study estimates in Figures 3.1 and 3.2 when the treatment is defined by whether any married woman and whether the wife of the household head in a given household was exposed to HSAA, respectively.

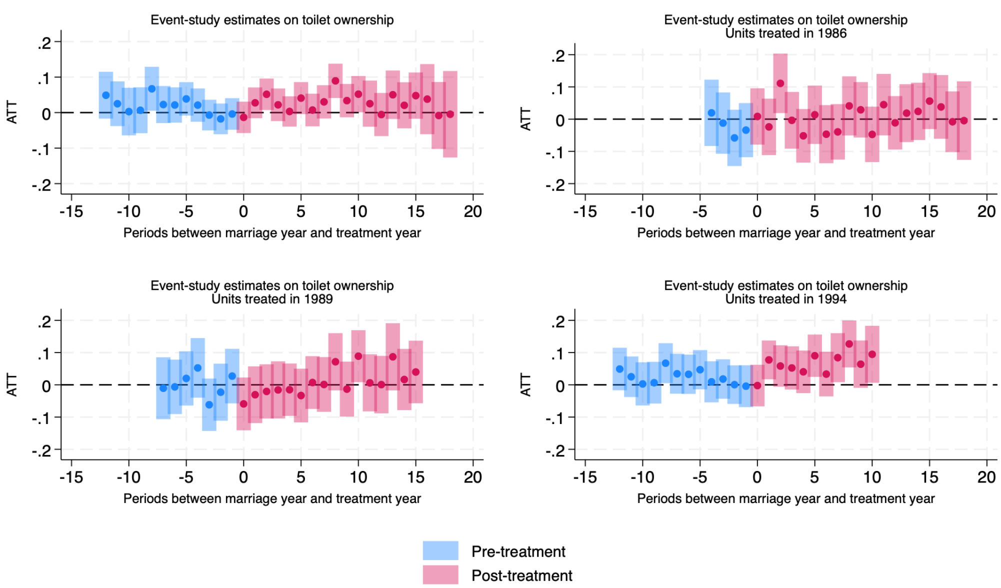
The event study figures show that there are no statistical differences between the treated and untreated states in the likelihood of the presence of a toilet in the households where women were married in years before the policy were implemented. This supports our conditional parallel trends assumption—in the absence of the policy, the evolution of toilet presence in households in treated states would have evolved in parallel to those in untreated

Figure 3.1: Event study on toilet ownership (Any married woman)



Notes: This figure plots event study estimates of the impact of the policy on the likelihood of woman's marital household having a toilet, by comparing marriage cohorts over time and using never-treated states as comparison groups.

Figure 3.2: Event study on toilet ownership (Wife of household head)



Notes: This figure plots event study estimates of the impact of the policy on the likelihood of woman's marital household having a toilet, by comparing marriage cohorts over time and using never-treated states as comparison groups.

states. Furthermore, for the event study, we take into account long differences to address any concerns surrounding pre-trends (Roth, 2013).

In the post-treatment periods, the event study plots show upward trends in toilet adoption for cohorts that got married at least 4-5 years after policy adoption in the late adopting states of Maharashtra and Karnataka.[20] Consistent with the results on the heterogenous treatment effects across groups we find no evidence of dynamic treatment effects in the early adopting states of Andhra Pradesh and Tamil Nadu.

3.7 Mechanisms

In order to understand the mechanisms behind these results, the natural question to ask is: Did the policy increase the likelihood of inheritance for treated women? If it did, then that would be a possible explanation behind these findings. However, existing work on HSA documents mixed evidence of the first order effects of the policy. Using data on land ownership from REDS (2006), Deininger, Goyal and Nagarajan (2013) specifically study whether the amendment was effective in the states of Maharashtra and Karnataka and find an increase in land ownership rates for treated women. On the contrary, using data from REDS (1999), Roy (2015) studies the impact of the policy for all the amendment states and finds no evidence of an increase in inheritance for treated women.[21] Since NFHS does not contain data on land ownership, this prevents us from examining the first order impact of the policy on land ownership by women in our context. However, even if the policy did not have the desired impact of improving access to inheritance rights for treated women, there exists consistent evidence that the policy led to an increase in alternate forms of parental

[20]Even though we observe a jump in the dynamic effects one year after policy adoption in states of Maharashtra and Karnataka, the upward trend in the policy effects in latter cohorts in consistent across event study estimates obtained by whether the treatment is defined by whether any married woman and whether the wife of the household head in a given household was exposed to HSAA, respectively.

[21]Another study by Agarwal, Anthwal and Mahesh (2021) makes use of longitudinal land ownership data collected by ICRISAT (years range from 2010-2014) and although they don't estimate the causal impact of the HSAA, their takeaway is that the policy was unsuccessful in improving women's inheritance rights.

investments in treated females, which can also be important in determining their later life wellbeing and socioeconomic outcomes.

Roy (2015) finds that the absence of these first order effects was because parents tried to circumvent the HSA by "gifting" away daughters' share of the inheritance to their brothers, but at the same time, also compensating daughters by increasing investment in their education, an alternate source of transfer.[22] This finding of compensating behavior on the part of parents and an increase in years of educational attainment for treated women under the HSAA has been widely documented (Deininger, Goyal and Nagarajan, 2013; Roy, 2015; Bose and Das, 2021; Ajefu et al., 2022). Evidence also suggests that the HSAA led to an increase in women's intra-household bargaining power (Deininger et al., 2019; Mookerjee, 2019; Bose and Das, 2021; Ajefu et al., 2022), and higher labor market participation (Heath and Tan, 2014). Studies that find an increase in bargaining power of women typically attribute their findings to higher years of educational attainment and unearned income through asset/land transfers under HSA. However, these papers focus on homogenous treatment effects in a setting of staggered policy implementation.

Our data allows us to test for two mechanisms that might drive the results on toilet ownership: we look at the women's years of educational attainment and their intra-household bargaining power within the marital household. We use the same estimation strategy as before, and also allow for heterogeneous and dynamic treatment effects to understand the timing of these mechanisms and whether they align with our main results.

Increased education can increase toilet coverage through the channel of increased awareness on the importance health and sanitation, or by increasing the chances of challenging pre-existing religious norms favoring open defecation. Improvements in intra-household bargaining power can also increase toilet coverage if women, who have higher innate preferences for toilets for reasons of protecting their privacy and dignity, lacked the intra-household say to construct them prior to policy. Another point worth noting here, is that as a response

[22] The author finds that this was possible due to the intestate nature of HSA, under which the rules of property division applied only in the absence of a will.

Table 3.3: Impact of HSAA on women's years of educational attainment (Any married woman)

	(1) Never treated	(2) Not yet treated
Aggregate ATT (GAverage)	0.421**	0.418**
	(0.175)	(0.175)
ATT of units treated in 1986	0.545*	0.522
	(0.326)	(0.322)
ATT of units treated in 1989	-0.228	-0.218
	(0.341)	(0.337)
ATT of units treated in 1994	0.646**	0.646**
	(0.254)	(0.254)
Observations	42,765	42,765

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

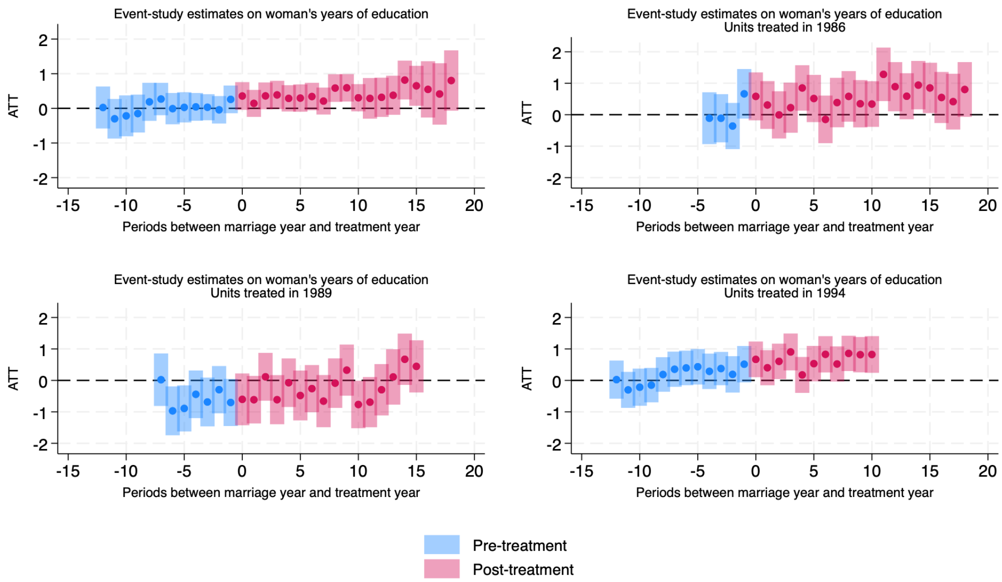
Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). We present estimates using never-treated as comparison. Standard errors are computed using wild cluster bootstrap at the state level.

to the policy, women might either choose to marry into households that had a toilet to begin with, or they might choose to construct it post marriage. Our data doesn't contain information on the year in which households get the toilet, and only contains information on whether the household owns a toilet in the survey year (2005, in our case), hence we are not able to test this directly. However, we do an indirect test by conducting our analysis with one of the earlier waves of NFHS (survey year 1992) where household level toilet ownership is measured in the year 1992, and find no evidence of the impact of the policy, suggesting that it was probably after marriage that treated women advocated for building a toilet.

3.7.1 Years of educational attainment

We report the estimates of heterogenous treatment effects of the HSAA in Table 3.3 defining treatment using retrospective exposure of any married woman to the HSAA. Consistent with our main results, we find that exposure to HSAA causes an increase in the years of educational attainment predominantly in the states that passed the amendment in 1994 by 0.64-0.79 years (over a base of 5 years) and is statistically significant at the 95% confidence level. These impacts in the late-adopting states are strong enough to drive an overall average

Figure 3.3: Event study on years of Education (Any married woman)



Notes: This figure plots event study estimates of the impact of policy on women's years of educational attainment, by comparing marriage cohorts over time and using never-treated states as comparison groups.

treatment effect of the HSAA on years of educational attainment. Consistent with our main results, we find little to no effect of the amendment on years of education in other states. Appendix Table C.2 shows similar results by using exposure of household head's wife as the treatment.

Allowing for dynamic treatment effects, we plot the corresponding event study estimates in Figure 3.3 which corroborate the results described in the previous paragraph. Here too we find an upward trend in education attainment for cohorts who married at least 4-5 years after the HSAA implementation in the later-adopting states suggesting that the policy primarily affected cohorts that were relatively young at the time of policy implementation in Maharashtra and Karnataka. This finding is similar to Roy (2015) and Deininger, Goyal and Nagarajan (2013), but we provide an additional insight that this result is primarily concentrated in the late adopting states with little to no effect in the early adopting states. Appendix Figure C.1 shows similar patterns when treatment is defined for household head's wife.

Table 3.4: Impact of HSAA on women's intra-household bargaining power (Any married woman)

	(1) Joint	(2) Decision-making	(3) Mobility	(4) Financial	(5) Low IPV
Aggregate ATT (GAverage)	0.125***	0.166***	0.0865*	0.00748	-0.0136
	(0.0469)	(0.0467)	(0.0442)	(0.0459)	(0.0490)
ATT of units treated in 1986	0.00766	0.0734	-0.0417	-0.0704	-0.0442
	(0.101)	(0.100)	(0.0918)	(0.0811)	(0.107)
ATT of units treated in 1989	0.225***	0.342***	0.193***	-0.0293	-0.0638
	(0.0823)	(0.0886)	(0.0715)	(0.0886)	(0.0999)
ATT of units treated in 1994	0.132**	0.132**	0.0974	0.0579	0.0214
	(0.0667)	(0.0648)	(0.0651)	(0.0682)	(0.0656)
Observations	42,765	42,765	42,765	42,765	42,765

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). We present estimates using never-treated as comparison. Standard errors are computed using wild cluster bootstrap at the state level.

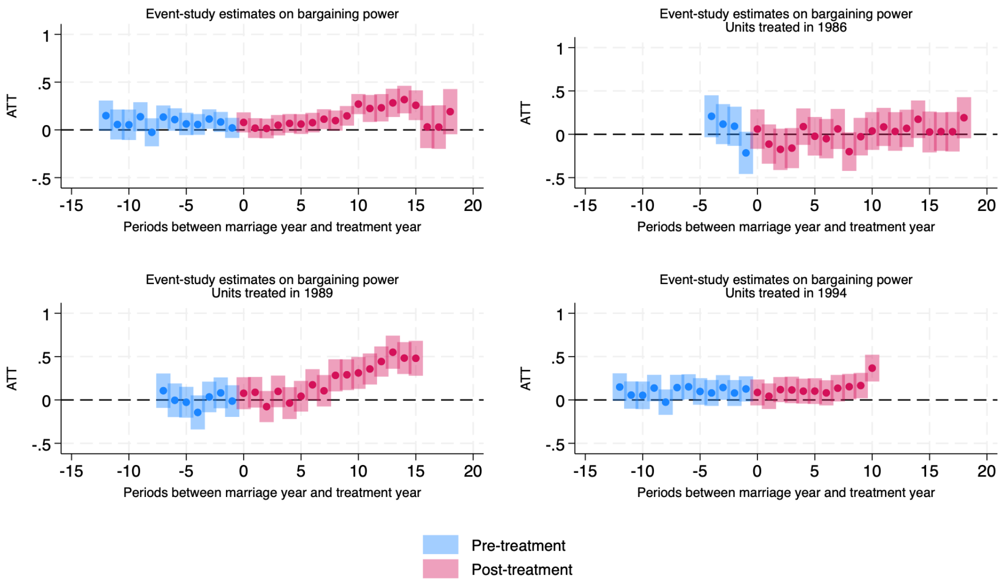
3.7.2 Intra-household bargaining power

We create indices of women's bargaining power using PCA, for four broad categories: household decision making, mobility, financial stability, and intimate partner violence.[23] We combine individual survey questions to create PCA-based indices on each of these four categories, coding each survey question as 1 to denote higher empowerment, and 0 otherwise. Then we use PCA to create an overall bargaining index by combining these four indices, and standardize it to create z-scores. The questions and data from NFHS that we use to create the four indices and their summary statistics are the same as used in Biswas, Das and Sarkhel (2024).[24]

[23]The household decision-making index is constructed by making use of the following survey questions: indicators for whether the woman makes decisions about her health care, major household purchases, purchases for daily household needs, and visiting family and relatives. Mobility index is constructed by using the following survey questions: indicators for whether the woman is allowed to go to the market, health facility, and places outside her village. Financial stability is constructed by using survey questions: indicators for whether the woman has any money that she alone decides to spend, and whether she has a bank account. Index for intimate partner violence is constructed using the following survey questions: indicators for whether husband is justified in hitting or beating his wife if she goes outside the house without telling him, neglecting the children, arguing with husband, refusing to have sex with husband, and not cooking food properly.

[24]Note that Biswas, Das and Sarkhel (2024) use these data to construct a proportion variable as the outcome. While our preferred measure is a z-score of the PCA-based index, our results are similar even when we use a proportion based outcome.

Figure 3.4: Event study on bargaining power (Any married woman)



Notes: This figure plots event study estimates of the impact of policy on women's intrahousehold bargaining power (joint index), by comparing marriage cohorts over time and using never-treated states as comparison groups.

As before we report the estimates of heterogenous treatment effects of the HSAA in Table 3.4. Column 1 shows the overall bargaining power index, while columns 2 - 5 denote the four categories that make the overall index. While the results in Table 3.4 show statistically significant coefficient estimates for units treated in 1994 (overall ATT increases by 0.12 SD units for treated women), the event study estimates for the same group in Figure 3.4 show weak evidence in support of this mechanism. We find that the increase in bargaining power for units treated in 1994 is driven by two marital cohorts that got married many years after the policy came into effect, which makes it a relatively weaker mechanism behind the observed gain in toilet ownership for this group. At the same time, for the group that gets treated in 1989, while there appears to be a substantial increase in bargaining power following HSA, however, that does not translate into higher toilet ownership rates for this group.

One would think that intra-household decision-making power of women might play a significant role in advocating for construction of a household toilet, however our results don't necessarily convey this idea. Figure 3.4 shows that even though women's bargaining power increased substantially in the state of Andhra Pradesh that passed the amendment in the year 1989, however there was no impact on toilet ownership rates for the state. This could be due to two reasons: first, our results show that the increase in bargaining-power in Andhra Pradesh was not accompanied by an increase in education levels. If education plays a key role in improving awareness regarding sanitation, then higher bargaining power within the household may not translate into advocating for building a toilet. Second, if we look at Appendix Figure C.2, where treatment is defined using exposure of the household head's wife, we see no impacts on her bargaining power whatsover, suggesting that the observed increase in bargaining power in Andhra Pradesh (Figure 3.4) can be entirely attributed to the daughters-in-law of the treated households. This further suggests that if women's bargaining power is indeed an important determinant in the construction of a household toilet, it is probably the bargaining power of the household matriarch, i.e, the wife of the household head, which plays a more critical role in important household purchases compared to the bargaining power of daughter-in-laws of the household.[25]

3.8 Robustness and potential concerns

In this section we list potential concerns which could threaten the identification of our parameter estimates. We provide evidence to show that our results are robust to these concerns and present additional robustness exercises.

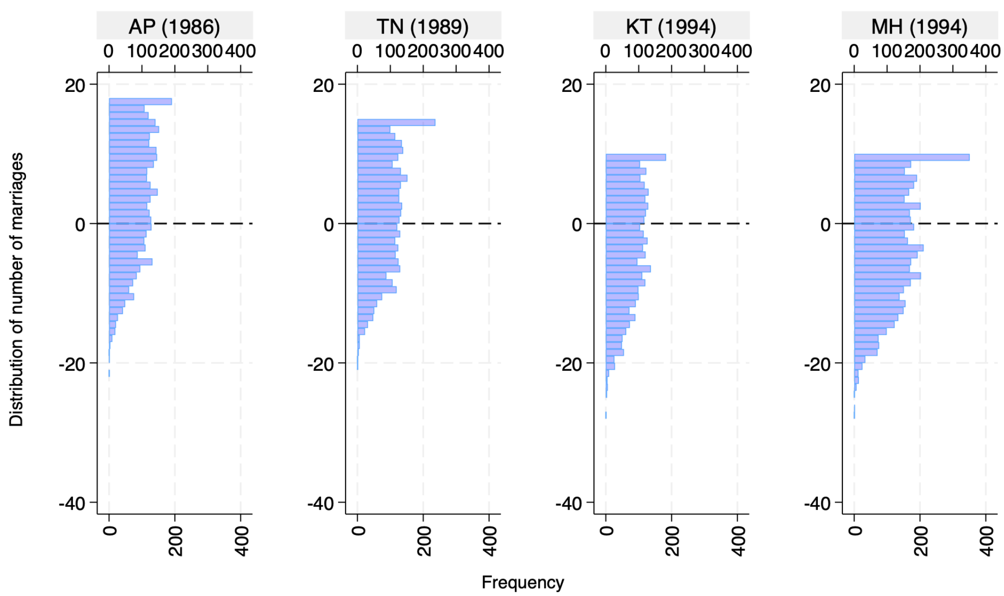
3.8.1 Endogenous selection into or out of policy

There are two concerns revolving around potential selection. On the one hand, if parents have a strong preference to endow the family inheritance to their sons relative to their daughters, they might respond by marrying off their daughters before the state level amendments. If

[25] Gupta, Ksoll and Maertens (2021), Anukriti et al. (2020) and Anukriti et al. (2022) show that in the Indian context, relationships between mother-in-laws and daughters-in-law are often characterized by power dynamics, and daughter-in-laws often lack the intra-household bargaining power to assert their preferences within the household.

this were the case, then such individuals are endogenously self-selecting out of the policy. On the other hand, gender-progressive families or individuals could potentially delay their marriage in order to be eligible for increased inheritance in anticipation of the policy. If this were the case, then this would lead such individuals to self-select into the treatment group and will not lead to clean comparisons in the difference-in-differences framework. Such self-selection patterns would be evident in the data by examining the distribution of year of marriage and age at marriage in the data. To address these concerns we plot these two variables in Fig 3.5 and Fig 3.6 but don't find any patterns of systemic jumps in marriages around the year of policy implementation. This tells us that it is unlikely that there was substantial self-selection into or out of the policy.

Figure 3.5: Distribution of marriages over time

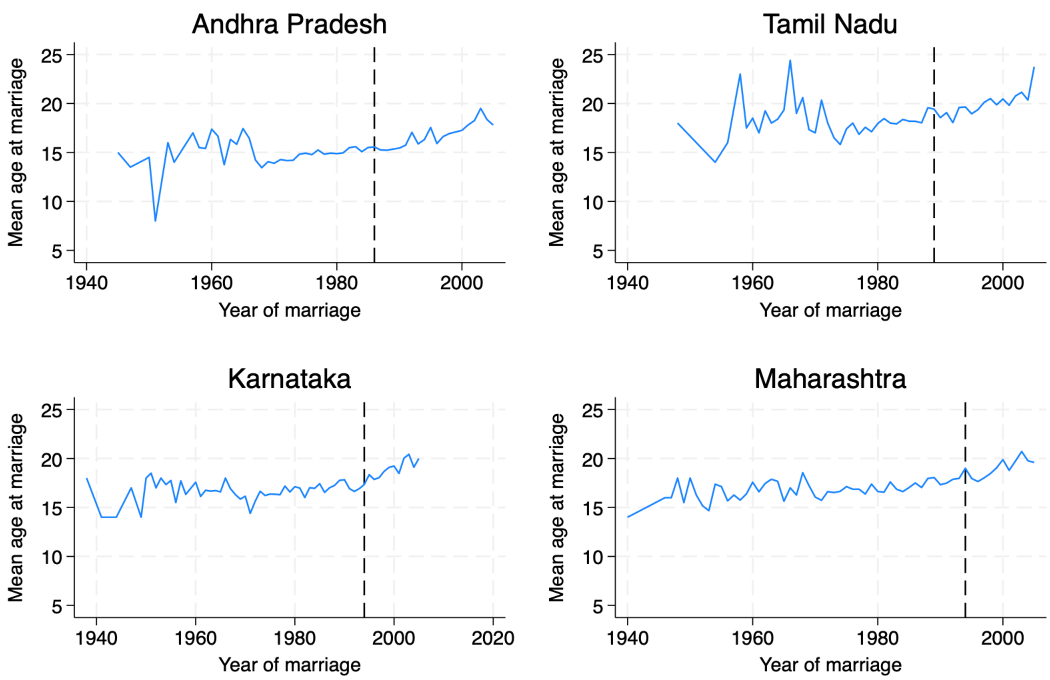


Notes: This figure plots the frequency distribution of the marriages by the difference in the year of marriage relative to the year of the state specific policy implementation. AP: Andhra Pradesh, TN: Tamil Nadu, KT: Karnataka and MH: Maharashtra.

3.8.2 Total Sanitation Campaign

Due to the dismal condition of sanitation in India in late 1990s, the Government of India introduced the Total Sanitation Campaign (TSC) in 1999. The TSC focused on increasing awareness about sanitation; however, it was not very successful in getting households to construct toilets (WSP, 2011). It was replaced by the Nirmal Bharat Abhiyaan in 2009, which provided subsidy payments for toilet construction to households below the poverty line. For the purpose of our identification, we assume that a national level policy like the TSC did not have differential effects across states in any given year.

Figure 3.6: Average age at marriage over time



Notes: The figure plots the average age at marriage of females over years. The dotted line in each sub-graph points on the x-axis the year of policy implementation. The spikes in the earlier years of marriage stem from very small sample sizes.

3.8.3 Post marital change in religion

We do not have data on females who have changed their religion, post-marriage. Failing to take this into account could result in biased estimates as religion is one of the criteria determining whether a woman benefited under the amendment. However, this is not much of a concern as inter-religious marriages are a rare occurrence in India. Das et al. (2011) provides evidence that only about 2.1% of marriages in India are inter-religious, citing social stigma as one of the biggest hindrances. Roy (2015) in her analyses of the effect of HSA on female education, finds only 3% of marriages to be inter-religious. Furthermore, the occurrence of inter-caste marriages within a religion is also rare. For example, Banerjee et al. (2013) show evidence of strong preference of marrying within the caste, to the extent that individuals are willing to trade off qualities like having a masters degree and no education. Thus, not being able to observe the above events is unlikely to change the results.

3.9 Conclusion

Existing research shows that females disproportionately benefit from having access to private toilets. However the lack of health and sanitation awareness, longstanding religious and cultural norms in favor of open defecation along with a lack of intra-household decision making power of females are key deterrents to take-up of toilets. In this paper, we examine the link between household sanitation and policies that empower females, being the first to do so to the best of our knowledge. We estimate the causal impact of the amendments to the Hindu Succession Act which were intended to improve female inheritance rights in India, on the presence of a toilet in treated women's marital households. We use a difference-in-differences framework with staggered adoption allowing for dynamic and heterogeneous treatment effects to estimate the impact of the reform on the presence of a toilet in the household. Given that one of the eligibility conditions is not observed in our, and in most datasets, we show that with modest assumptions, the estimate serves as a lower bound on the parameter of interest. We find a positive and statistically significant impact of improved inheritance rights for females on the presence of a toilet in their marital household. Allowing for heterogenous treatment effects, we find that the results are primarily concentrated in late adopting states. Allowing for dynamic treatment effects we find that the effect of the policy was the strongest for women who were relatively young at the time of policy amendment in their state and were married 4-5 years after the policy amendment. Our paper highlights that policies that empower women can serve as an unexpected yet effective strategy for improving sanitation coverage in regions struggling with open-defecation issues.

Appendix A

Appendix: Chapter 1

A.1 Figures

Figure A.1: Timeline of events

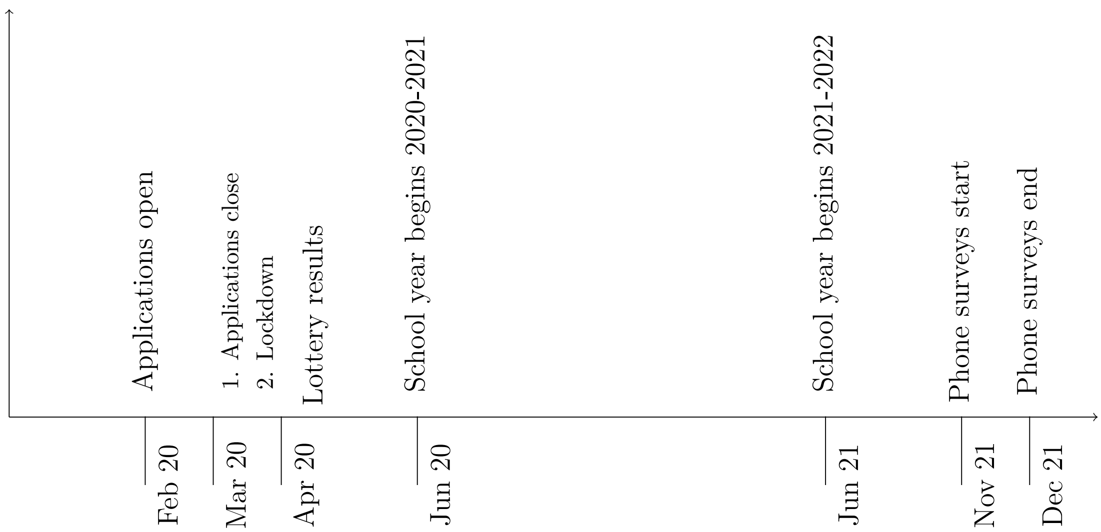
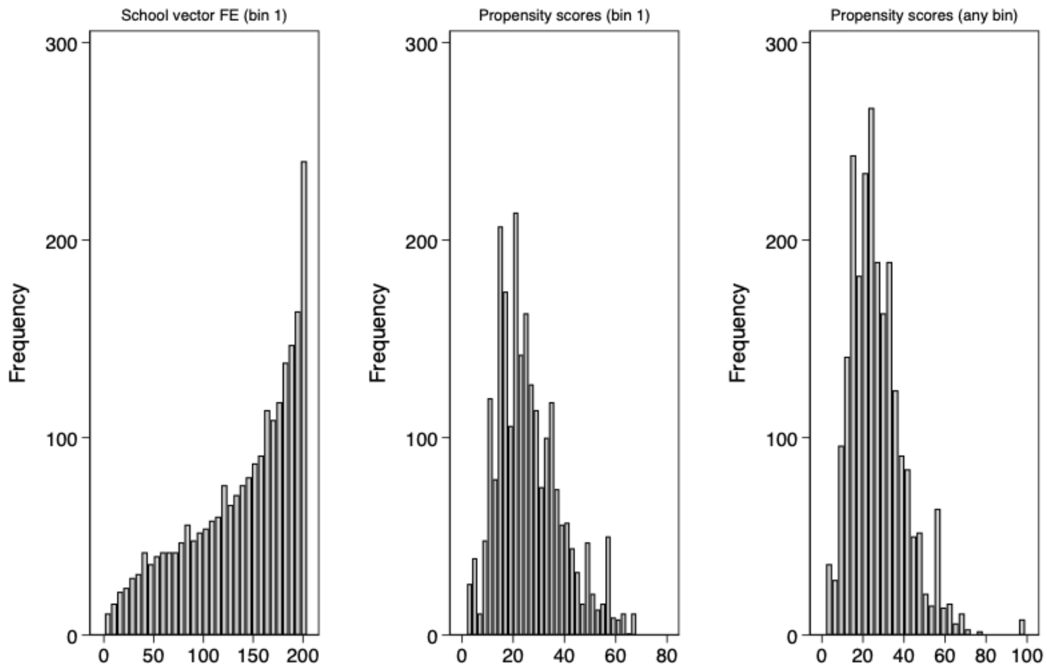


Figure A.2: Distribution of school vector fixed effects and ex-ante propensity scores of winning



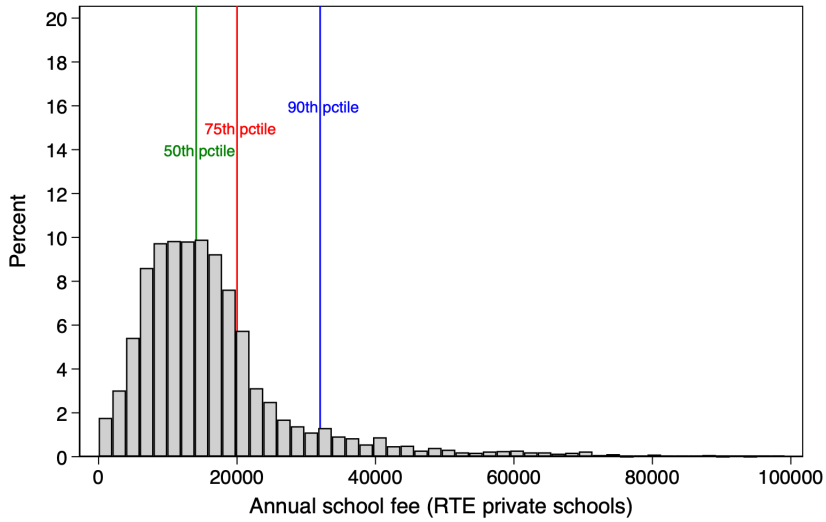
Notes: This is a histogram showing the distribution of school vector fixed effects (chosen in bin 1), and the simulated ex-ante propensity scores of winning the school lottery (in distance bin 1, and in any distance bin). The sample comprises surveyed applicants. In the sample of surveyed applicants, there are a total of 204 unique school vectors that are chosen in bin 1. A total of 193 school vectors out of these 204 vectors contribute to the identifying within-vector variation, i.e., they have at least one winner and at least one loser within bin 1. Distribution of simulated ex-ante propensity score bins in distance bin 1, and in any distance bin is also plotted. Here the propensity score bins are 0.01 interval wide.

Figure A.3: English and Math questions asked during phone based assessments

Measure	Question and answers
1	If someone asks you "What is your name" and "What is your gender" then what would you reply? (phrase in quotes is said in English, the rest is said in Hindi.)
	correct; incorrect
2	Can you recite the letters of the English Alphabet?
	correct; incorrect
3	Can you tell me the spelling of "BOAT"?
	correct; incorrect
4	Can you tell me the spelling of "SWIM"?
	correct; incorrect
5	If you have 9 chocolates, and you get 1 more chocolate, how many chocolates will you have in total?
	correct (answer = 10); incorrect (answer ≠ 10)
6	If you have 22 chocolates, and you get 38 more chocolates, how many chocolates will you have in total?
	correct (answer = 60); incorrect (answer ≠ 60)
7	If you have 20 chocolates, and you give 4 chocolates to your friend, how many chocolates will you be left with?
	correct (answer = 16); incorrect (answer ≠ 16)
8	If you have 45 chocolates, and you give 26 chocolates to your friend, how many chocolates will you be left with?
	correct (answer = 19); incorrect (answer ≠ 19)
9	Can you tell me the number of "tens" and ones in the number 96?
	correct(answer = 9 tens and 6 ones)

Notes: This table shows the list of questions asked to children during phone-based assessments. For all questions, the question was said in Hindi, but the key phrases/numbers were said in English. For example, the following things were said in English - the phrase in quotes "What is your name" and "What is your gender" (for question 1); English Alphabet (for question 2); the words "Boat" and "Swim" (for question 3, 4); numbers like 9 chocolates, 20 chocolates etc. (for questions 5-9).

Figure A.4: Distribution of annual school fee for RTE private schools



Notes: This histogram shows the distribution of annual school fee (in INR) for all the RTE private schools in Maharashtra. The data comes from the official website of the State Department of Education, Maharashtra.

Figure A.5: Factor loadings from first component of PCA

Proportion of functional toilets (boys)	0.1804
Proportion of functional toilets (girls)	0.1517
School building is privately owned	0.309
School has pucca (brick and mortar) boundary walls	0.2265
School has library	0.0813
School has playground	0.1261
School has computer lab	0.2303
School has Internet	0.4024
Laptops per pupil	0.0751
Desktops per pupil	0.2322
Printers per pupil	0.144
Digiboard per pupil	0.173
English medium	0.3638
Proportion of teachers with undergraduate college degree or higher	0.1255
Proportion of teachers with Bachelors in Education degree or higher	0.3086
Proportion of regular teachers	0.2413
Proportion of teachers below age 55	0.0207
Proportion of teachers not involved in non-teaching tasks	0.2952
Proportion of children belonging to general caste	0.2465

Notes: This shows the factor loadings on each of the variable that is used in the construction of the school quality index using principal component analysis. The first component explains 18% variation in the data.

Figure A.6: Tabulating eliteness across fee and PCA index measure

School is Elite (PCA > 50th pctile)

School is Elite (Fee > 50th pctile)

	NO	YES	MISSING	Total
NO	44	36	0	80
YES	33	132	0	165
MISSING	6	7	8	21
Total	83	175	8	266

School is Elite (PCA > 75th pctile)

School is Elite (Fee > 75th pctile)

	NO	YES	MISSING	Total
NO	103	18	0	121
YES	60	64	0	124
MISSING	6	7	8	21
Total	169	89	8	266

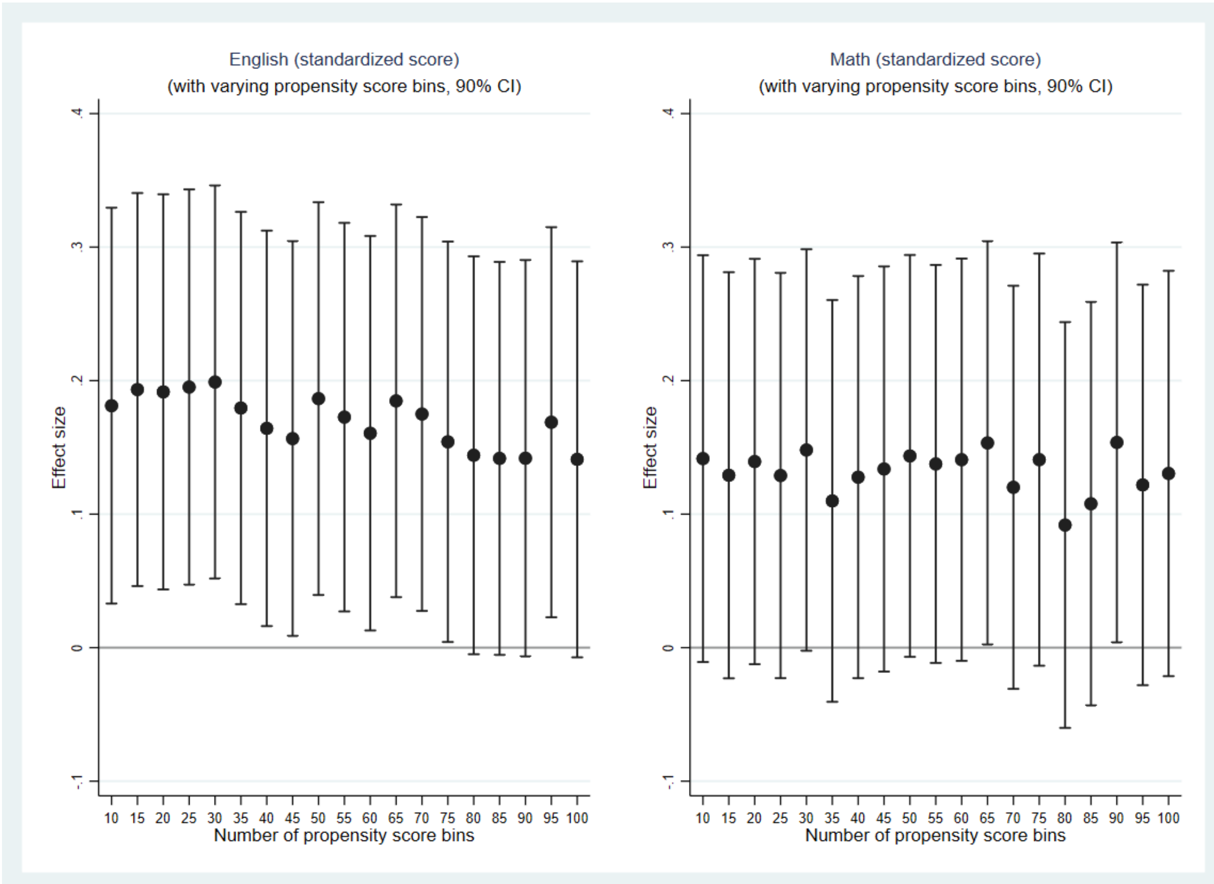
School is Elite (PCA > 90th pctile)

School is Elite (Fee > 90th pctile)

	NO	YES	MISSING	Total
NO	159	14	0	173
YES	56	16	0	72
MISSING	13	0	8	21
Total	228	30	8	266

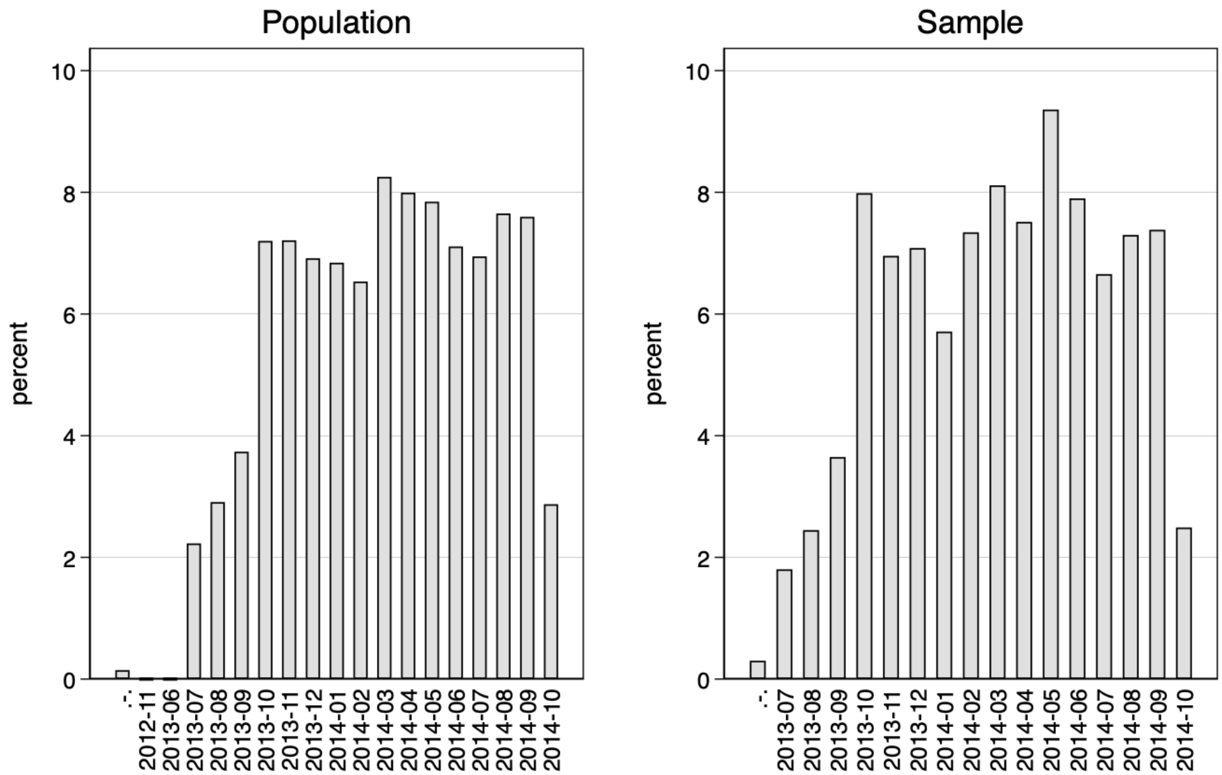
Notes: This provides a cross-tabulation of schools chosen by lottery winners, based on whether the school is categorized as elite or budget as per the PCA index measure and the school fee measure.

Figure A.7: Robustness: LATE of being a RTE quota student on phone-based assessments



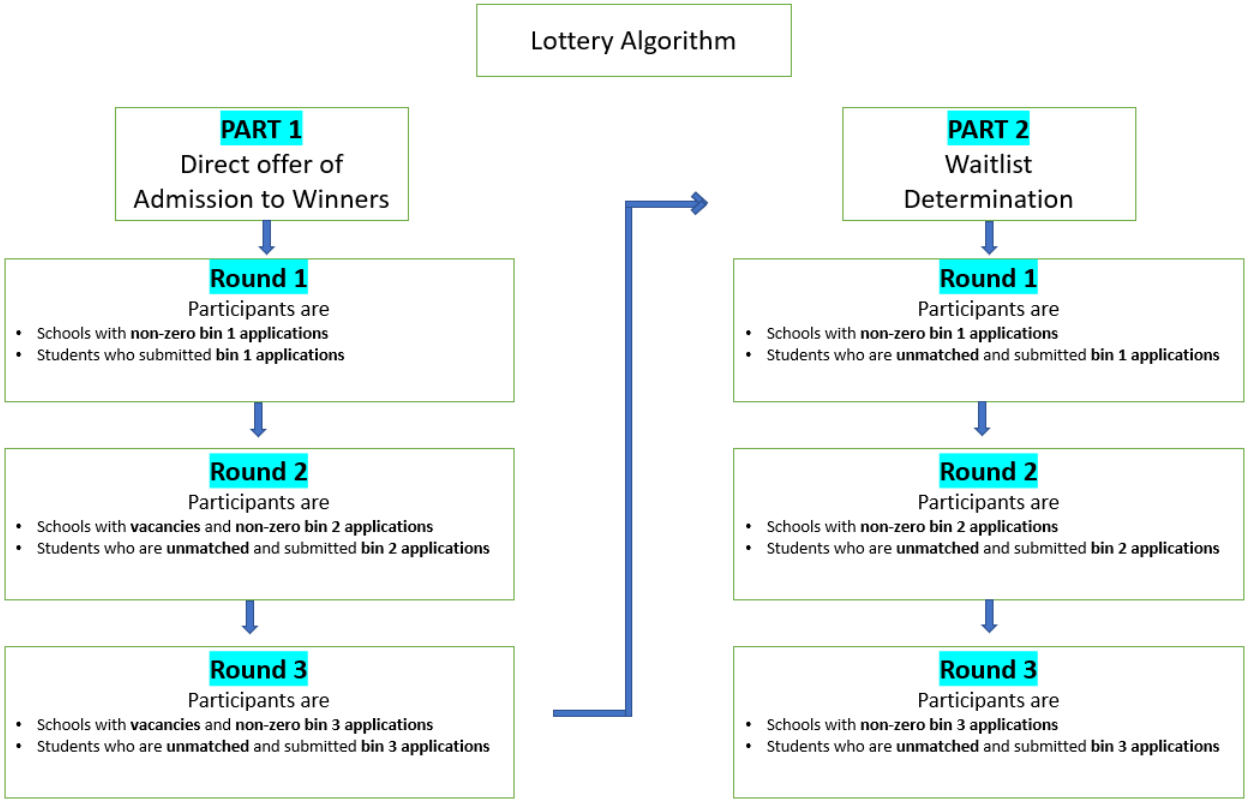
Notes: This figure plots the LATE of enrolling as an RTE student on children's performance in English and Math. It shows how the LATE changes as the number of bins of ex-ante propensity scores of winning are increased. The within comparisons become stricter as the number of propensity score bins are increased. The number of propensity score bins vary from 10, 15, 20, ..., 100. This utilizes the within variation resulting from comparison of treated and control students who have a similar ex-ante propensity of winning.

Figure A.8: Robustness: Histogram of birth year and month



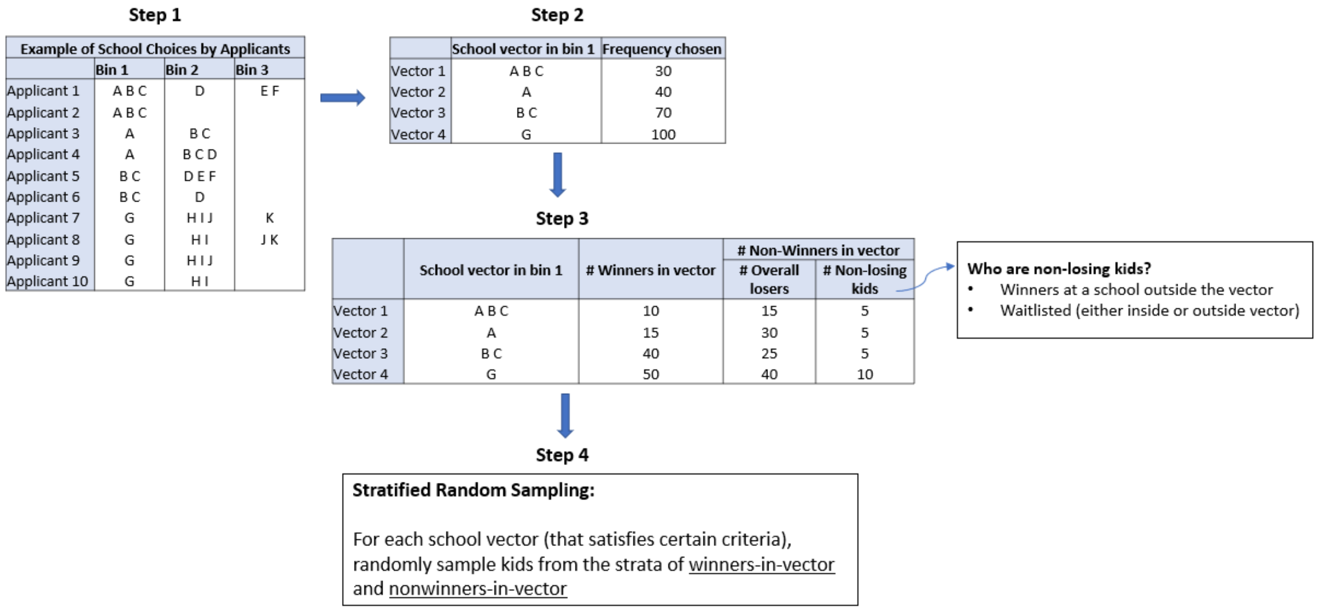
Notes: This figure shows the histogram of birth year-months for applicants to grade 1 private school lotteries under RTE policy in the 2020-21 school year. The left panel shows the distribution for the population and the right panel shows it for the sample. Some missing values exist. Birth year-months given by 2012-11 and 2013-06 are pertaining to only disability quota applicants and only appear in the population histogram. Disability quota is chosen very rarely and constitutes only 0.6% of the applications in the population. My sample does not contain any disability quota applicants. The majority of applications for grade 1 in 2020-21 school year can be seen as coming from those born in July 2013 and October 2014. Among these, applicants born between July 2014 and October 2014 are age-eligible to re-apply for grade 1 in the following year i.e., during the 2021-22 RTE lotteries. In one of my robustness checks, I remove these applicants who were still age-eligible to re-apply for the RTE lotteries in the 2021-22 school year, and find that my results are robust to removing them.

Figure A.9: Schematic flowchart explaining the lottery algorithm



Notes: This flowchart explains the lottery algorithm which the state of Maharashtra uses to allocate schools to applicants under the RTE 25% reservation policy at private schools. The allocation mechanism is a two part process, starting with determining the winners (Part 1, as shown in the left panel), followed by determining the waitlisted candidates (Part 2, as shown in the right panel).

Figure A.10: Schematic flowchart explaining sampling strategy



A.2 Tables

Table A.1: Summary statistics

Variable	N	Mean	SD	Min	Max
Characteristics of applicants in Phone Survey					
Winner (distance bin 1)	2,329	0.44	0.50	0	1
Winner (any distance bin)	2,329	0.45	0.50	0	1
Waitlisted (any distance bin)	2,329	0.26	0.44	0	1
Loser	2,329	0.29	0.46	0	1
Age	2,329	7.62	0.33	7.05	8.34
Male	2,329	0.55	0.50	0	1
Number of schools chosen (RTE application)	2,329	4.86	2.89	1	10
Applied under low income quota	2,329	0.28	0.45	0	1
Schooling details for applicants					
Academic year: 2020-21					
School enrollment	2,329	0.89	0.31	0	1
School provides instruction	2,083	0.89	0.31	0	1
Academic year: 2021-22					
School enrollment	2,329	0.97	0.18	0	1
School provides instruction	2,255	0.98	0.14	0	1
School is Private	2,255	0.88	0.32	0	1
School is English medium	2,255	0.94	0.25	0	1
Instructional days at school	2,255	5.30	1.42	0	7
Number of subjects taught	2,107	5.93	1.85	1	12
Monetary investments in applicants					
Any educational expense (in the past year)	2,227	0.93	0.26	0	1
Annual educational expenses (INR; in the past year)	2,227	3,514	3,234	0	24,000
Time investments in applicants					
Child gets help with homework in the household	2,329	0.93	0.26	0	1
Hours of household help with homework (hours per week)	2,329	9.50	5.91	0	49
Time use of applicants					
Attending school (hours per week)	2,329	12	7.98	0	36
Attending tuition (hours per week)	2,329	4.67	6.10	0	21
Doing homework (hours per day)	2,329	1.40	0.73	0	3.30
Playing (hours per day)	2,328	2.45	1.18	0	6
Watching Television (hours per day)	2,322	1.10	0.91	0	4
Helping with household chores (hours per day)	2,329	0.39	0.40	0	2
Performance on phone assessments by applicants					
English score (standardized)	695	-0.00	1.00	-1.56	1.92
Math score (standardized)	695	-0.00	1.00	-1.81	1.91
Parental education					
Mother's education > primary	2,329	0.62	0.49	0	1
Fathers's education > primary	2,329	0.54	0.50	0	1
Household characteristics					
Number of household members	2,329	5.14	2.10	2	20
Number of siblings of applicant child	2,329	0.88	0.57	0	5
General Caste	2,329	0.26	0.44	0	1
Scheduled Caste	2,329	0.25	0.43	0	1
Scheduled Tribe	2,329	0.04	0.19	0	1
Other Backward Class (OBC)	2,329	0.46	0.50	0	1
Hindu	2,329	0.81	0.39	0	1
Muslim	2,329	0.09	0.29	0	1
Buddhist	2,329	0.09	0.29	0	1
Other religion	2,329	0.01	0.09	0	1
Household SES index (PCA)	2,329	0.00	1.21	-2.51	6.23
Annual household earnings (INR 1000)	2,001	180	132	2.40	1,200

Notes: This table shows the summary statistics of survey participants who comprise the sample. Most of the data in this table comes from phone-survey data conducted during the months of Nov-Dec 2021 (18 months after RTE results came out). Characteristics of applicants, religion, and caste information comes from the administrative data of RTE applications. Some of the variables are conditional on other variables, such as indicator of whether school provides instruction, and other variables under schooling details, are conditioned on school enrollment. Monetary investments are asked for the past year i.e., 2020-21, and includes expenses on child's education on stationary, books etc. (excluding school fee). Time investments by parents and household members is calculated by asking about time spent helping child with educational activities on a typical day of the week in the past week (along with number of days). Applicants' time use is calculated by asking about time spent on each activity on a typical day in the past week, and additionally, number of days per week for variables that measure weekly hours. English and Math scores are standardized - the English assessment had four questions, the Math assessment had five questions. Household SES index is created using Principal Components Analysis using data on asset ownership of television, air conditioner, two-wheeler, and four-wheeler.

Table A.2: Balance in baseline characteristics

Variable	(1) Non winners (any bin)	(2) Winners (any bin)	(3) Difference ((2)-(1))
Age of applicant (as on 1st Nov 2021)	7.622	7.611	-0.014
	(0.326)	(0.329)	(0.014)
Male	0.545	0.560	0.012
	(0.498)	(0.497)	(0.021)
Schools chosen overall (RTE application)	4.935	4.759	-0.094
	(2.905)	(2.877)	(0.113)
Applied under low income quota	0.288	0.275	-0.015
	(0.453)	(0.447)	(0.019)
Mother's education > primary	0.871	0.890	0.022
	(0.336)	(0.313)	(0.014)
Father's education > primary	0.822	0.849	0.030*
	(0.383)	(0.359)	(0.016)
Number of household members	5.130	5.149	0.028
	(2.119)	(2.078)	(0.089)
Number of siblings of applicant	0.881	0.874	-0.011
	(0.586)	(0.544)	(0.024)
General Caste	0.261	0.251	-0.012
	(0.439)	(0.434)	(0.018)
Scheduled Caste	0.259	0.233	-0.023
	(0.439)	(0.423)	(0.018)
Scheduled Tribe	0.038	0.034	-0.006
	(0.191)	(0.181)	(0.008)
Other Backward Classes	0.442	0.482	0.042**
	(0.497)	(0.500)	(0.021)
Hindu	0.795	0.823	0.030*
	(0.404)	(0.382)	(0.017)
Muslim	0.097	0.086	-0.011
	(0.296)	(0.280)	(0.012)
Buddhist	0.098	0.086	-0.014
	(0.298)	(0.280)	(0.012)
Other religion	0.010	0.006	-0.004
	(0.100)	(0.076)	(0.004)
Household SES index (PCA)	0.057	-0.071	-0.117**
	(1.250)	(1.145)	(0.051)
Annual income from survey (INR 1000)	189.750	163.519	-25.531***
	(127.564)	(105.862)	(5.405)
Observations	1,291	1,038	2,329

Notes: This table shows the balance in baseline characteristics across non-winning and winning applicants (in any bin). The differences in column (3) control for the fixed effects of ex-ante propensity of winning the lottery in any bin such that the comparisons across winners and losers are for ex-ante similar applicants. Columns (1) and (2) contain the mean and standard deviation of the variables for non-winners and winners. Column (3) contains the coefficient in front of the dummy of being a winner from the regression of the outcome variable (displayed in the rows) on the dummy of winning, after controlling for the ex-ante propensity of winning in any bin (propensity score bins are 0.01 wide). Column (3) shows standard errors in parentheses.

Table A.3: Attrition: participation in the phone survey

Panel A: Sample includes everyone who was ever called for phone surveys

	(1) Participation, Survey = 0	(2) Participation, Survey = 1	(3) Difference ((2)-(1))
Winner (any distance bin)	0.394	0.446	0.057***
	(0.489)	(0.497)	(0.016)
Age of applicant (as on 1st Nov 2021)	7.639	7.617	-0.022**
	(0.335)	(0.328)	(0.011)
Male	0.523	0.547	0.027*
	(0.500)	(0.498)	(0.016)
Schools chosen overall (RTE application)	4.753	4.857	0.055
	(2.928)	(2.893)	(0.065)
Applied under low income quota	0.296	0.282	0.010
	(0.457)	(0.450)	(0.013)
General Caste	0.265	0.257	0.016
	(0.441)	(0.437)	(0.013)
Scheduled Caste	0.250	0.248	0.002
	(0.433)	(0.432)	(0.013)
Scheduled Tribe	0.033	0.036	-0.002
	(0.179)	(0.186)	(0.006)
Other Backward Classes	0.452	0.459	-0.016
	(0.498)	(0.498)	(0.015)
Hindu	0.794	0.807	0.020*
	(0.404)	(0.395)	(0.012)
Muslim	0.102	0.092	-0.008
	(0.303)	(0.289)	(0.008)
Buddhist	0.093	0.093	-0.008
	(0.290)	(0.290)	(0.009)
Other religion	0.011	0.008	-0.003
	(0.104)	(0.090)	(0.003)
Observations	1,930	2,329	4,259

Panel B: Sample includes everyone who agreed to participate in surveys

	(1) Participation, Phone Assessments = 0	(2) Participation, Phone Assessments = 1	(3) Difference ((2)-(1))
Winner (any distance bin)	0.437	0.466	0.026
	(0.496)	(0.499)	(0.024)
Age of applicant (as on 1st Nov 2021)	7.610	7.635	0.029*
	(0.332)	(0.316)	(0.016)
Male	0.575	0.479	-0.099***
	(0.494)	(0.500)	(0.024)
Schools chosen overall (RTE application)	4.765	5.072	0.147
	(2.881)	(2.913)	(0.094)
Applied under low income quota	0.285	0.275	-0.015
	(0.452)	(0.447)	(0.019)
General Caste	0.260	0.249	-0.012
	(0.439)	(0.433)	(0.019)
Scheduled Caste	0.242	0.260	0.002
	(0.429)	(0.439)	(0.020)
Scheduled Tribe	0.035	0.039	-0.001
	(0.184)	(0.193)	(0.009)
Other Backward Classes	0.463	0.452	0.011
	(0.499)	(0.498)	(0.022)
Hindu	0.810	0.800	-0.014
	(0.392)	(0.400)	(0.018)
Muslim	0.091	0.095	0.005
	(0.287)	(0.293)	(0.012)
Buddhist	0.091	0.098	0.011
	(0.287)	(0.297)	(0.013)
Other religion	0.009	0.007	-0.003
	(0.092)	(0.085)	(0.004)
Observations	1,634	695	2,329

Table A.4: Characteristics of lottery compliers, always- and never-takers in Maharashtra's RTE

	Compliers			
Variable	Untreated (1)	Treated (2)	Always-takers (3)	Never-takers (4)
Male	0.545	0.560	0.561	0.540
	(0.019)	(0.017)	(0.010)	(0.010)
Low income quota applicant	0.295	0.276	0.300	0.248
	(0.017)	(0.015)	(0.009)	(0.009)
Caste quota applicant	0.704	0.723	0.699	0.751
	(0.017)	(0.015)	(0.009)	(0.009)
General caste	0.274	0.258	0.263	0.201
	(0.017)	(0.015)	(0.009)	(0.008)
Scheduled Caste	0.245	0.216	0.299	0.302
	(0.017)	(0.014)	(0.009)	(0.009)
Scheduled tribe	0.043	0.035	0.019	0.016
	(0.007)	(0.006)	(0.002)	(0.002)
Other caste	0.437	0.489	0.417	0.480
	(0.019)	(0.017)	(0.010)	(0.010)
Hindu	0.796	0.833	0.796	0.773
	(0.015)	(0.013)	(0.008)	(0.009)
Muslim	0.096	0.082	0.106	0.094
	(0.011)	(0.009)	(0.006)	(0.006)
Buddhist	0.094	0.077	0.096	0.123
	(0.011)	(0.009)	(0.006)	(0.007)
Mother education > primary	0.868	0.897	0.903	0.849
	(0.013)	(0.010)	(0.006)	(0.007)
Father education > primary	0.822	0.859	0.752	0.846
	(0.015)	(0.013)	(0.009)	(0.008)
Mother works	0.233	0.212	0.230	0.260
	(0.016)	(0.014)	(0.009)	(0.009)
Father works	0.943	0.946	0.949	0.967
	(0.008)	(0.007)	(0.004)	(0.003)
Share of observations	.81		.07	.12

Notes: This table reports the estimates of average baseline characteristics of compliers, always-takers, and never-takers among lottery applicants to private schools under Maharashtra's RTE quotas. Means are computed from 2SLS and OLS regressions that control for lottery risk set indicators (or,ex-ante propensity scores of winning the lottery), as described in Abadie (2002) (see Appendix Section A.4.1 for details on implementation). Robust standard errors in parentheses.

Table A.5: Counterfactual densities for Maharashtra's RTE Compliers

Destiny	RTE Private school
	Z=0
	(1)
Fee-paying student at RTE Private school	0.564
	(0.019)
Fee-paying student at Non-RTE Private school	0.112
	(0.013)
Government school	0.191
	(0.015)
Out-of-school	0.052
	(0.009)
At school (but can't match school)	0.087
	(0.011)
Pscores of winning	Yes

Notes: This table reports the share of untreated (Z=0) compliers enrolled at particular fallback school types among applicants to Maharashtra's RTE private school lotteries. Means are computed from 2SLS regressions that control for the ex-ante propensity scores of winning the lottery, as described in Abadie (2002). I describe the implementation of this in Appendix Section A.4.1. Among lottery losers, there are some children for whom the school name and the official school code could not be matched with the administrative data on the population of schools. Thus, for these children the school sector – private, government, or out-of-school – is missing. It is for this reason that the counterfactual destinies don't add up to one. Robust standard errors in parentheses.

Table A.6: LATE of being a RTE quota student on school quality index

	Joint index	Infrastructure index	Digital index	Teacher index	Peer SES index
	(1)	(2)	(3)	(4)	(5)
Enrolled as RTE student	0.613***	0.329***	0.434***	0.411***	0.219***
	(0.050)	(0.053)	(0.052)	(0.052)	(0.048)
First stage F-stat	3,608.10	3,608.10	3,608.10	3,608.10	3,608.10
Outcome mean	-0.00	0.00	0.00	0.00	-0.00
Control mean	-0.27	-0.15	-0.15	-0.21	-0.11
Observations	2,086	2,086	2,086	2,086	2,086
R^2	0.20	0.10	0.14	0.14	0.28
Pscores of winning	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on school quality, when the instrument is winning the lottery in any bin. Controls include - sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.7: LATE of being a RTE quota student on subjects taught

	Math	English	Marathi	Hindi	Science	Environmental studies
	(1)	(2)	(3)	(4)	(5)	(6)
Enrolled as RTE student	0.017	0.028**	-0.001	0.112***	0.021	0.113***
	(0.014)	(0.014)	(0.019)	(0.022)	(0.020)	(0.026)
First stage F-stat	3,877.71	3,877.71	3,877.71	3,877.71	3,877.71	3,877.71
Outcome mean	0.92	0.93	0.84	0.78	0.18	0.53
Control mean	0.91	0.92	0.84	0.73	0.17	0.48
Observations	2,255	2,255	2,255	2,255	2,255	2,255
R^2	0.07	0.06	0.05	0.09	0.05	0.10
Pscores of winning	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Computers	General knowledge	Art/craft	Music	Dance	Physical education
	(7)	(8)	(9)	(10)	(11)	(12)
Enrolled as RTE student	0.136***	0.096***	0.108***	0.067***	0.049***	-0.000
	(0.024)	(0.024)	(0.024)	(0.013)	(0.012)	(0.011)
First stage F-stat	3,877.71	3,877.71	3,877.71	3,877.71	3,877.71	3,877.71
Outcome mean	0.29	0.32	0.31	0.07	0.05	0.95
Control mean	0.23	0.27	0.27	0.04	0.03	0.96
Observations	2,255	2,255	2,255	2,255	2,255	2,255
R^2	0.08	0.08	0.09	0.08	0.08	0.06
Pscores of winning	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on the subjects taught at school when the instrument is winning the lottery in any bin. Outcomes measure the indicator of whether school teaches a particular subject. Controls include - sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.8: Time investments in children by household members: Extensive Margin

	Mother	Father	Grandparents	Siblings	Uncle/Aunt	Neighbors
	(1)	(2)	(3)	(4)	(5)	(6)
Enrolled as RTE student	0.056***	0.026	0.010	-0.014	0.000	-0.005
	(0.020)	(0.023)	(0.006)	(0.016)	(0.010)	(0.004)
First stage F-stat	3,916.37	3,916.37	3,916.37	3,916.37	3,916.37	3,916.37
Outcome mean	0.81	0.28	0.01	0.11	0.04	0.01
Control mean	0.79	0.26	0.01	0.11	0.04	0.01
Observations	2,329	2,329	2,329	2,329	2,329	2,329
R^2	0.12	0.10	0.04	0.05	0.07	0.02
Pscores of winning (any bin)	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending RTE private schools as a quota student on the indicator of whether a specific household member helps the child with educational activities. The outcome variables capture the extensive margin of whether child gets any help from mom, dad, and grandparents. Controls include - sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.9: First stage of winning the RTE lottery (in bin 1) on enrollment as a RTE quota student

	Enrolled as RTE student	
	(1)	(2)
Instrument = Winning lottery in Bin 1	0.790***	0.787***
	(0.013)	(0.013)
Outcome mean	0.44	0.44
Control mean	0.09	0.09
Observations	2,329	2,329
R^2	0.66	0.64
School vector FE (bin 1)	Yes	No
Pscores of winning (bin 1)	No	Yes
Controls	Yes	Yes

Notes: This table shows the first stage effects of winning the RTE private school lottery in distance bin 1, on enrollment as an RTE quota student in a private school. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Column (1) controls for the fixed effects of school vector chosen in bin 1, and Column (2) controls for the ex-ante propensity of winning the lottery in bin 1. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.10: LATE of being a RTE quota student (in bin 1) on enrollment

	Enrollment (2021-22)		Enrollment (2021-22)		Grade 2 and above (2021-22)	
	(1)	(2)	(3)	(4)	(5)	(6)
Enrolled as RTE student	0.130***	0.140***	0.045***	0.047***	0.186***	0.194***
	(0.015)	(0.016)	(0.009)	(0.009)	(0.017)	(0.017)
First stage F-stat	3,589.78	3,751.65	3,617.60	3,776.54	3,611.65	3,772.53
Outcome mean	0.89	0.89	0.97	0.97	0.86	0.86
Control mean	0.84	0.84	0.94	0.94	0.78	0.78
Observations	2,328	2,328	2,328	2,328	2,327	2,327
R^2	0.20	0.11	0.13	0.08	0.20	0.15
School vector FE (bin 1)	Yes	No	Yes	No	Yes	No
Pscores of winning (bin 1)	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on children's enrollment when the instrument is winning the lottery in bin 1. The outcomes measure the indicator of school enrollment in the two academic years. Controls include - sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Odd numbered columns control for the fixed effects of school vector chosen in bin 1, and even numbered columns control for the ex-ante propensity of winning the lottery in bin 1. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.11: LATE of being a RTE quota student (in bin 1) on test scores

	Test score (standardized)			
	English		Math	
	(1)	(2)	(3)	(4)
Enrolled as RTE student	0.169**	0.178*	0.093	0.135
	(0.085)	(0.091)	(0.090)	(0.093)
First stage F-stat	947.07	1,194.62	947.07	1,194.62
Outcome mean	-0.00	-0.00	-0.00	-0.00
Control mean	-0.10	-0.10	-0.09	-0.09
Observations	695	695	695	695
R^2	0.41	0.16	0.33	0.11
School vector FE (bin 1)	Yes	No	Yes	No
Pscores of winning (bin 1)	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on children's performance on phone-based assessments when the instrument is winning the lottery in bin 1. Outcomes measure children's standardized test scores on English and Math. Controls include - sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Odd numbered columns control for the fixed effects of school vector chosen in bin 1, and even numbered columns control for the ex-ante propensity of winning the lottery in bin 1. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.12: School characteristics in Elite and Budget schools

	Fee			PCA		
Variable	Budget (1)	Elite (2)	Difference (3)	Budget (4)	Elite (5)	Difference (6)
Proportion functional toilets (boys)	1.000	0.988	0.000	0.994	0.996	-0.000
	(0.000)	(0.101)	(0.000)	(0.079)	(0.030)	(0.000)
Proportion functional toilets (girls)	0.995	0.988	0.012	0.990	0.996	0.009
	(0.059)	(0.101)	(0.015)	(0.095)	(0.030)	(0.015)
School building is privately owned	0.437	0.602	0.172	0.356	0.870	0.464***
	(0.498)	(0.492)	(0.108)	(0.480)	(0.339)	(0.100)
School building has pucca boundary	0.778	0.971	0.009	0.812	0.986	0.142**
	(0.417)	(0.169)	(0.068)	(0.392)	(0.120)	(0.067)
School has library	0.937	0.981	0.016	0.938	1.000	0.100*
	(0.245)	(0.139)	(0.058)	(0.243)	(0.000)	(0.058)
School has playground	0.905	0.981	0.051	0.925	0.971	0.062
	(0.295)	(0.139)	(0.062)	(0.264)	(0.169)	(0.063)
School has computer lab	0.143	0.350	0.125	0.156	0.420	0.265***
	(0.351)	(0.479)	(0.078)	(0.364)	(0.497)	(0.076)
School has internet	0.849	1.000	0.139**	0.881	1.000	0.161**
	(0.359)	(0.000)	(0.065)	(0.325)	(0.000)	(0.065)
Laptops per pupil	0.003	0.005	0.004	0.003	0.005	0.005*
	(0.005)	(0.013)	(0.003)	(0.007)	(0.014)	(0.003)
Desktops per pupil	0.025	0.037	0.007	0.023	0.046	0.025***
	(0.028)	(0.034)	(0.008)	(0.026)	(0.037)	(0.008)
Printers per pupil	0.004	0.005	0.001	0.004	0.005	0.002*
	(0.003)	(0.004)	(0.001)	(0.003)	(0.005)	(0.001)
Digiboards per pupil	0.002	0.008	0.007***	0.002	0.011	0.008***
	(0.005)	(0.011)	(0.002)	(0.005)	(0.012)	(0.002)
School is English medium	0.849	1.000	0.169**	0.881	1.000	0.166**
	(0.359)	(0.000)	(0.071)	(0.325)	(0.000)	(0.072)
Prop. of teachers trained in computer	0.449	0.628	0.255***	0.511	0.571	-0.013
	(0.386)	(0.367)	(0.088)	(0.386)	(0.388)	(0.092)
Prop. of teachers who are graduates	0.736	0.859	0.074	0.759	0.867	0.079
	(0.292)	(0.197)	(0.055)	(0.281)	(0.186)	(0.055)
Prop. of teachers with Bachelors in Education	0.421	0.682	0.200***	0.462	0.716	0.213***
	(0.282)	(0.192)	(0.057)	(0.281)	(0.167)	(0.058)
Prop. of full time teachers	0.673	0.759	0.005	0.725	0.680	-0.157*
	(0.417)	(0.362)	(0.095)	(0.394)	(0.397)	(0.094)
Prop. of contract teachers	0.316	0.232	0.011	0.268	0.303	0.149
	(0.415)	(0.364)	(0.093)	(0.392)	(0.403)	(0.093)
Prop. of part-time teachers	0.011	0.009	-0.016	0.007	0.017	0.008
	(0.065)	(0.027)	(0.015)	(0.047)	(0.060)	(0.015)
Prop. of teachers $<$ 55 years	0.953	0.965	0.009	0.951	0.975	0.049**
	(0.107)	(0.068)	(0.021)	(0.106)	(0.038)	(0.020)
Prop. of teachers not involved in non-teaching tasks	0.859	0.919	-0.015	0.855	0.959	0.068
	(0.295)	(0.234)	(0.059)	(0.302)	(0.159)	(0.060)
Teachers per pupil	0.037	0.040	0.010	0.037	0.042	0.010
	(0.029)	(0.019)	(0.006)	(0.027)	(0.021)	(0.006)
Prop. of general caste category students	0.273	0.527	0.097***	0.334	0.512	0.075***
	(0.283)	(0.271)	(0.025)	(0.296)	(0.290)	(0.026)
Observations	126	103	229	160	69	229

Notes: This table shows the balance in school characteristics for elite and budget schools, where eliteness is defined using the two measures: school fee and PCA index. Schools lying above the 75^{th} percentile value in the distribution of fee and PCA index of all the private schools in the state are classified as elite schools, and classified as budget, otherwise. The sample comprises schools being attended by lottery winners. Columns (1), (2), (3), and (4) show the mean and standard-deviations of the characteristics for budget and elite schools based on the two quality measures. Columns (3) and (4) contain the coefficient on the indicator of "school is elite" from the regression of the outcome variable (dispalyed in rows) on the indicator of school being elite, after controlling for the geography fixed effects at the village level (standard errors in parentheses).

Table A.13: Correlation between fee-based school eliteness and PCA-based school eliteness

	School Fee (log)
	(1)
School quality index (standardized)	0.272***
	(0.015)
Outcome mean	9.51
Observations	4,019
R^2	0.51
Village FE	Yes

Notes: This table shows the regression of the log of school fee on the school quality index on the population of RTE schools for whom there is non-missing data on school fee.

Table A.14: LATE of attending an elite schools as a RTE quota student on subjects taught

	Math	English	Marathi	Hindi	Science	Envt studies	Comp-uters	General knowledge	Art/ craft	Music	Dance	Phys ed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Elite based on Fee												
RTE student	0.014	0.029	-0.043	-0.020	-0.111**	0.033	0.054	0.241***	0.145**	0.138***	0.085**	-0.013
Elite school	(0.036)	(0.033)	(0.051)	(0.054)	(0.055)	(0.070)	(0.067)	(0.067)	(0.069)	(0.041)	(0.036)	(0.032)
F-stat	1,090	1,090	1,090	1,090	1,090	1,090	1,090	1,090	1,090	1,090	1,090	1,090
Outcome mean	0.93	0.94	0.85	0.83	0.18	0.59	0.35	0.36	0.36	0.10	0.08	0.95
Control mean	0.92	0.93	0.84	0.79	0.18	0.52	0.26	0.28	0.32	0.03	0.02	0.95
Observations	965	965	965	965	965	965	965	965	965	965	965	965
R^2	0.08	0.08	0.08	0.08	0.07	0.08	0.12	0.12	0.06	0.17	0.16	0.10
Pscores	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel B: Elite based on PCA												
RTE student	0.039	0.025	0.072	0.015	-0.057	0.153*	0.117	0.182**	0.041	0.121**	0.126***	0.004
Elite school	(0.042)	(0.038)	(0.058)	(0.061)	(0.063)	(0.078)	(0.076)	(0.076)	(0.077)	(0.047)	(0.042)	(0.035)
F-stat	1,099	1,099	1,099	1,099	1,099	1,099	1,099	1,099	1,099	1,099	1,099	1,099
Outcome mean	0.93	0.94	0.84	0.82	0.19	0.58	0.35	0.36	0.36	0.10	0.07	0.95
Control mean	0.92	0.93	0.85	0.80	0.20	0.53	0.31	0.33	0.36	0.07	0.06	0.96
Observations	1,011	1,011	1,011	1,011	1,011	1,011	1,011	1,011	1,011	1,011	1,011	1,011
R^2	0.07	0.08	0.07	0.07	0.05	0.08	0.08	0.09	0.07	0.09	0.05	0.10
Pscores	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on the subjects taught at child's school. The sample is restricted to lottery winners. Envt studies refers to Environment studies, and Phy Ed refers to Physical Education. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.15: First stage (varying the elite cutoff)

	RTE student at Elite school	
	Elite (PCA)	Elite (Fee)
	(1)	(2)
Panel A: Eliteness defined at 50^{th} pctile		
Won RTE lottery at Elite school	0.878***	0.881***
	(0.032)	(0.031)
Outcome mean	0.67	0.65
Control mean	0.00	0.00
Observations	1,019	973
R^2	0.70	0.72
Pscores of winning at elite	Yes	Yes
Controls	Yes	Yes
Avg quality (Elite=1)	3.79	37143.46
Avg quality (Elite=0)	1.56	10292.29

	RTE student at Elite school	
	Elite (PCA)	Elite (Fee)
	(1)	(2)
Panel B: Eliteness defined at 90^{th} pctile		
Won RTE lottery at Elite school	0.865***	0.850***
	(0.022)	(0.026)
Outcome mean	0.10	0.31
Control mean	0.00	0.00
Observations	1,019	973
R^2	0.89	0.86
Pscores of winning at elite	Yes	Yes
Controls	Yes	Yes
Avg quality (Elite=1)	4.96	54733.34
Avg quality (Elite=0)	2.71	16090.5

Notes: This table shows the first stage effects of winning the RTE private school lottery at an elite school on enrollment at an elite school as a quota student. Here, I present the results with two different percentile cutoffs of eliteness - at 50^{th} and 90^{th} percentile in panel A and B, respectively. The sample is restricted to lottery winners. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.16: LATE of attending elite schools on performance on tests (varying the elite cutoff)

	English	Math	English	Math
	Elite (PCA)		Elite (Fee)	
	(1)	(2)	(3)	(4)
Panel A: Eliteness defined at 50^{th} pctile				
RTE student at Elite school	0.083	0.256	0.189	0.241
	(0.252)	(0.250)	(0.217)	(0.220)
First stage F-stat	229.44	229.44	333.36	333.36
Outcome mean	0.04	0.06	0.05	0.06
Control mean	-0.12	-0.08	-0.18	-0.19
Observations	318	318	303	303
R^2	0.12	0.15	0.16	0.18
Pscores of winning at elite	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

	English	Math	English	Math
	Elite (PCA)		Elite (Fee)	
	(1)	(2)	(3)	(4)
Panel B: Eliteness defined at 90^{th} pctile				
RTE student at Elite school	0.219	0.510	0.024	-0.094
	(0.421)	(0.422)	(0.308)	(0.309)
First stage F-stat	1,821.94	1,821.94	567.59	567.59
Outcome mean	0.04	0.06	0.05	0.06
Control mean	0.02	0.02	-0.07	-0.02
Observations	318	318	303	303
R^2	0.11	0.12	0.14	0.17
Pscores of winning at elite	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on children's performance on phone based assessments. The results correspond to the 50^{th} and 90^{th} percentile cutoffs of eliteness in panel A and B, respectively. The sample is restricted to lottery winners. As before, the number of observations is smaller here because the phone-based assessment on English and Math is available only for a subsample of lottery winners. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.17: LATE of attending elite schools on school instruction (varying the elite cutoff)

	Synchronous classes (online)	Recordings shared (audio/video)	Text-based activity plans (WhatsApp/SMS)	Synchronous classes (online)	Recordings shared (audio/video)	Text-based activity plans (WhatsApp/SMS)
	Elite (PCA)			Elite (Fee)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Eliteness defined at 50^{th} pctile						
Quota student at Elite school	0.195***	-0.021	-0.162**	0.244***	-0.110**	-0.194***
	(0.050)	(0.046)	(0.068)	(0.046)	(0.045)	(0.066)
First stage F-stat	788.37	788.37	788.37	864.10	864.10	864.10
Outcome mean	0.83	0.13	0.53	0.83	0.13	0.53
Control mean	0.73	0.13	0.55	0.65	0.18	0.60
Observations	1,005	1,005	1,005	959	959	959
R^2	0.16	0.07	0.09	0.23	0.07	0.11
Pscores of winning at elite	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

	Synchronous classes (online)	Recordings shared (audio/video)	Text-based activity plans (WhatsApp/SMS)	Synchronous classes (online)	Recordings shared (audio/video)	Text-based activity plans (WhatsApp/SMS)
	Elite (PCA)			Elite (Fee)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: Eliteness defined at 90^{th} pctile						
Quota student at Elite school	0.108	-0.005	0.257*	0.196***	-0.018	-0.128
	(0.103)	(0.092)	(0.139)	(0.063)	(0.058)	(0.083)
First stage F-stat	1,453.58	1,453.58	1,453.58	1,326.33	1,326.33	1,326.33
Outcome mean	0.83	0.13	0.53	0.83	0.13	0.53
Control mean	0.81	0.13	0.52	0.77	0.15	0.58
Observations	1,005	1,005	1,005	959	959	959
R^2	0.08	0.04	0.04	0.11	0.06	0.11
Pscores of winning at elite	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on school's instruction modality. The results correspond to the 50^{th} and 90^{th} percentile cutoffs of eliteness in panel A and B, respectively. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.18: LATE of attending elite schools on children's time use
(varying the elite cutoff)

	School (hrs/week)	Tuition (after school) (hrs/week)	Homework (hrs/day)	School (hrs/week)	Tuition (after school) (hrs/week)	Homework (hrs/day)
	Elite (PCA)			Elite (Fee)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Eliteness defined at 50^{th} pctile						
Quota student at Elite school	2.874***	-0.405	0.091	3.272***	-0.825	-0.006
	(0.985)	(0.815)	(0.095)	(0.988)	(0.810)	(0.095)
First stage F-stat	725.95	725.95	725.95	750.79	750.79	750.79
Outcome mean	13.58	4.43	1.51	13.57	4.49	1.52
Control mean	12.60	4.30	1.45	12.18	4.49	1.45
Observations	1,019	1,019	1,019	973	973	973
R^2	0.13	0.09	0.06	0.13	0.10	0.06
Pscores of winning at elite	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

	School (hrs/week)	Tuition (after school) (hrs/week)	Homework (hrs/day)	School (hours/week)	Tuition (after school) (hrs/week)	Homework (hrs/day)
	Elite (PCA)			Elite (Fee)		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: Eliteness defined at 90^{th} pctile						
Quota student at Elite school	2.885	-2.674	0.312	0.528	0.040	-0.290**
	(2.053)	(1.650)	(0.194)	(1.304)	(1.047)	(0.124)
First stage F-stat	1,331.97	1,331.97	1,331.97	985.43	985.43	985.43
Outcome mean	13.58	4.43	1.51	13.57	4.49	1.52
Control mean	13.38	4.58	1.50	12.86	4.58	1.51
Observations	1,019	1,019	1,019	973	973	973
R^2	0.07	0.08	0.04	0.08	0.09	0.03
Pscores of winning at elite	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending elite RTE private schools as a quota student, on children's time use. The results correspond to the 50^{th} and 90^{th} percentile cutoffs of eliteness in panel A and B, respectively. Control variables include - sex and age of child, indicators for father's and mother's education being greater than the respective means, indicator of low income quota applicant, household's SES index, indicator of caste categories, and religion. Simulated ex-ante propensity scores of winning the lottery in any distance bin are controlled. Results are robust to increasing the number of propensity score bins. Robust standard errors in parentheses.

Table A.19: Robustness to Attrition using Inverse-Probability Reweighting

	English	Math
	(1)	(2)
Enrolled as RTE student	0.332***	0.293**
	(0.126)	(0.119)
Total observations	695	695
Treatment observations	324	324
Control observations	371	371

Notes: This table shows the results for the LATE of attending private schools as a quota student on children's test scores, by using inverse probability weighting to account for the differential probability of attrition or non-response based on baseline observables.

Table A.20: Robustness: Excluding young applicants to estimate LATE of being a quota student on test scores

	Test score (standardized)	
	English	Math
	(1)	(2)
Enrolled as RTE student	0.169*	0.180*
	(0.097)	(0.103)
First stage F-stat	1,054.22	1,054.22
Outcome mean	-0.01	-0.01
Control mean	-0.10	-0.11
Observations	590	590
R^2	0.14	0.08
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: This table shows the results for the LATE of attending private schools as a quota student on children's test scores, by excluding young applicants from the sample who are age-eligible to apply again under RTE in the year 2021-22.

Table A.21: Robustness: LATE of being a quota student on school instruction

	School provides instruction (2021-22)	Synchronous (online) (2021-22)	Recordings shared (audio/video) (2021-22)	Text based activity plans (WhatsApp/SMS) (2021-22)
	(1)	(2)	(3)	(4)
Enrolled as RTE student	0.020***	0.152***	-0.069***	-0.085***
	(0.005)	(0.019)	(0.015)	(0.025)
First stage F-stat	3,562.84	3,557.02	3,557.02	3,557.02
Outcome mean	0.99	0.80	0.10	0.57
Control mean	0.98	0.75	0.13	0.59
Observations	2,255	2,238	2,238	2,238
R^2	0.05	0.18	0.10	0.11
Pscores of winning	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports the estimated coefficient β^{LATE} from the 2SLS regression that estimates the local average treatment effect of attending a private school as a quota student on school instruction and school modality. This table uses a new variable which is generated such that it is unique at the school level, and captures a unique response to each school being attended in the sample. To do this, I recode the new variable equal to the value that is reported by the majority of the applicants (at least 50%) attending that school. For example, if more than 50% children attending school A say that school was providing instruction, then I code the variable to reflect that school A was providing instruction (for each child who is enrolled at that school, regardless of their original response). Column (1) looks at the dummy of whether school provides any instruction in the 2020-21 academic year, and columns (2), (3), and (4) look at the instructional modality offered by school. Thus, the outcome here is recoded such that there is a unique value associated with each school. Controls include - sex and age of child, indicators for father's and mother's education being greater than the mean, dummy of low income quota applicant, SES index, dummies of caste categories, and religion. Robust standard errors in parentheses.

Table A.22: Distribution of simulated ex-ante propensity scores of winning

Panel A: Population

Variable	N	10th pctile	25th pctile	50th pctile	75th pctile	90th pctile	95th pctile	99th pctile
NAGPUR								
For winners	6,330	0.20	0.31	0.51	0.80	1	1	1
For waitlisted	5,913	0.03	0.16	0.30	0.45	0.58	0.67	0.81
For losers	9,974	0	0	0.06	0.18	0.27	0.33	0.45
PUNE								
For winners	15,198	0.25	0.42	0.65	0.94	1	1	1
For waitlisted	13,606	0	0.12	0.30	0.47	0.62	0.71	0.86
For losers	13,385	0	0	0.01	0.15	0.26	0.32	0.43
THANE								
For winners	8,041	0.42	0.75	1.00	1	1	1	1
For waitlisted	3,756	0	0.09	0.28	0.47	0.64	0.77	0.93
For losers	1,392	0	0	0.00	0.17	0.27	0.32	0.42
MUMBAI								
For winners	4,721	0.33	0.55	0.94	1	1	1	1
For waitlisted	2,776	0	0.16	0.33	0.48	0.65	0.73	0.89
For losers	1,727	0	0	0.08	0.19	0.27	0.36	0.45

Panel B: Sample

Variable	N	10th pctile	25th pctile	50th pctile	75th pctile	90th pctile	95th pctile	99th pctile
NAGPUR								
For winners	584	0.14	0.20	0.27	0.35	0.48	0.55	0.62
For waitlisted	318	0.17	0.25	0.33	0.41	0.49	0.56	0.62
For losers	396	0.12	0.16	0.22	0.29	0.34	0.38	0.53
PUNE								
For winners	275	0.10	0.15	0.20	0.33	0.52	0.57	0.68
For waitlisted	154	0.12	0.16	0.22	0.38	0.56	0.58	0.68
For losers	228	0.08	0.11	0.16	0.21	0.26	0.30	0.38
THANE								
For winners	134	0.22	0.26	0.30	0.39	0.46	0.71	1
For waitlisted	108	0.22	0.25	0.34	0.40	0.44	0.52	0.64
For losers	43	0.21	0.22	0.28	0.31	0.39	0.44	0.45
MUMBAI								
For winners	45	0.21	0.26	0.36	0.42	0.56	0.64	1
For waitlisted	28	0.25	0.26	0.39	0.53	0.62	0.63	1
For losers	16	0.19	0.20	0.25	0.26	0.27	0.39	0.39

Notes: This table shows the distribution of simulated ex-ante propensity scores of winning under the lottery mechanism.

A.3 Lottery Algorithm, Sampling, and Simulation of Algorithm

A.3.1 Lottery algorithm

Here I explain the lottery algorithm that is implemented for RTE 25% quotas in the state of Maharashtra.

Part 1: Direct offer of admission to winners

1. Schools are arranged in the descending order of total applications received under the policy in the *previous* year. Based on this rank ordering, each school gets their turn to do the allotment of students in the *current* year.

2. There are three rounds in which the allotment happens. Each round corresponds to one of the three distance bins in which schools receive applications.

   Round 1- schools receiving applications in bin 1.

3. The first round comprises each school that received non-zero applications from students who applied to the school in distance bin 1 and allotment is done only for students who applied to these schools in bin 1.

4. The top school (as determined by the rank ordering of schools) allocates seats by lottery if the count of applications received in bin 1 > seats at the school. The school allocates seats to all bin 1 applicants without any lottery if the count of applications received in bin 1 $\leq$ seats. Within the bin, all applicants are treated equally and thus have the same ex-ante probability of being selected in the lottery.[1] All the applicants who are matched to this school are removed from the consideration set and only unmatched

[1] This mechanism satisfies the Equal Treatment of Equals (ETE) property following Abdulkadiroğlu, Angrist, Narita and Pathak (2017). ETE is said to satisfy when students with the same preferences and priorities have the same chance of getting allocated at any given school.

applicants are considered for further matching. The school is removed from further matching if it has exhausted all its vacancies.[2]

5. Revised bin-level demand is calculated for all the remaining schools. The previous step is repeated for the next school based on the rank ordering list of schools. The school conducts a lottery based admission if the revised demand by bin 1 applicants exceeds the number of vacancies at school. This process is iterated over all the schools, while maintaining the same initial rank ordering.

6. After the end of round 1, all applicants have been considered at all their bin 1 school choices and all schools have tried to allot any available seats by offering them to their respective bin 1 applicants.

Round 2- schools having applications from unalloted applicants in bin 2

7. Next is the second round. The second round comprises schools which have non-zero vacancies and have non-zero applications from those who applied here in bin 2, based on revised bin-level demand at the end of round 1. In this round, allotment is only done for applicants who (i) failed to get a seat in round 1 and had applied somewhere in bin 2, and (ii) applicants who only applied to bin 2 schools.

8. The allotment process is same as before. The top school (based on the same initial rank ordering of schools) allots seats by lottery if the count of revised applications in bin 2 $>$ seats. School allots seats to everyone who applied here in bin 2 without a lottery if the count of revised applications in bin 2 $\leq$ seats.

9. Revised bin level demand is calculated for all the remaining schools, and the previous step is iterated over all the remaining schools, following the same initial rank ordering of schools.

[2] If a school conducts a lottery to admit children in round 1 (i.e., for those who applied in the nearest distance bin), then this means that the school will not admit students who applied in the other two distance bins.

10. At the end of round 2, all applicants who were remaining to be matched after round 1 and were bin 2 applicants somewhere, plus applicants who only applied to schools in bin 2 have been considered at all their bin 2 school choices, conditional on the fact that these school still had seats to offer.

    Round 3- schools having applications from unalloted applicants in bin 3

11. Next is the third round. The idea is same as before. Schools which feature here are those that still have vacancies after rounds 1 and 2. Hence round 3 considers applicants who are (i) remaining to be matched after the end of round 2 and had applied somewhere in bin 3, and (ii) applicants who only applied to schools in bin 3.

12. The allotment process is same as before. The top school (based on the same initial rank ordering of schools) allots seats via lottery if the count of revised applications in bin 3 $>$ vacant seats. School allots seats to everyone who applied here in bin 3 without a lottery if the count of revised applications in bin 3 $\leq$ vacant seats.

13. This marks the end of direct offer of admissions to winners.

Part 2: Waitlist determination

Even after the previous steps described in Part 1, there are many applicants who are yet to be matched. These applicants are either waitlisted at a unique school or are rejected from all the schools. There are 3 rounds in which the waitlist determination happens. The process is exactly similar to Part 1 and is explained as follows:

1. Schools are arranged in the same initial rank ordering as before and take turns to do the allotment based on this rank ordering. The rule is that the maximum number of waitlisted students at a school is equal to the number of winners at the school (where the number of winners per school is established in Part 1).

    Round 1- schools having applications from unalloted applicants in bin 1

2. Round 1 comprises schools which have unmatched applications from those residing in bin 1 (these are applicants who did not get matched in Part 1).

3. The top school provides offers of waitlist by lottery if the count of unmatched applications in bin 1 > seats available under waitlist. Within the bin, all applicants are treated equally in the event of a lottery. Each matched applicant is assigned a waitlist priority at the school which determines the ordering in which they will be called for admission in the event that any winner at this school forgoes their seat.[3] All matched applicants are removed from the consideration set and the school is removed from any further matching if it has exhausted all its vacancies. Revised bin-level demand is calculated for all remaining schools. This process is iterated for all the remaining schools, following the same initial ranking.

   Round 2- schools having applications from unalloted applicants in bin 2

4. Round 2 comprises schools which have unmatched applications from those residing in bin 2 (these are applicants who did not get matched either in Part 1 or round 1 of waitlist). Similar as before, step 3 is iterated at each eligible school, taking into account unmatched applications received in bin 2.

   Round 3- schools having applications from unalloted applicants in bin 3

5. Round 3 marks the final round. This comprises schools which have unmatched applications from bin 3 students (these are applicants who did not get matched either in Part 1 or in round 1, and 2 of the waitlist determination). Step 3 is iterated at each eligible school, taking into account unmatched applications received in bin 3.

6. At the end of Round 3, there are still some applicants who are remaining to be matched anywhere. These are the applicants who are not selected anywhere and I refer to them as overall lottery losers.

[3]The waitlist priority assigned to applicants at each school is randomly generated.

A.3.2 Sampling strategy

1. I focus on the districts for whom I have the most complete administrative data (Mumbai, Nagpur, Pune, and Thane), and focusing on these districts I make a list of all unique combinations of schools chosen by applicants in the nearest distance bin (henceforth, bin 1). This gives me all the unique *school vectors* that were chosen in bin 1. By virtue of this, some school vectors consist of a single school, and some consist of multiple schools.

2. For each school vector chosen in bin 1, I compute the count of winners who win at any school in the vector (given by the sum of winners at each school in the vector) and count of non-winners who did not win at any school they listed in bin 1.

3. Thus, non-winners for a given school vector comprise applicants who might be: (a) winners at a school that was chosen in distance bin 2 or 3, (b) waitlisted at a school that was chosen in distance bin 1, 2, or 3, and (c) overall losers who lost their chance at each and every school that they listed in each distance bin.[4]

4. Next, I focus on those school vectors which meet the following criteria:

    (i). Count of winners in the vector is at least 4.

    (ii). Count of overall losers in the vector is at least 4.

    (iii). Share of overall losers (among non-winners) in the vector is at least 0.75.

    As an aside - The rule about count of winners and overall losers being at least 4, was imposed taking into account the possibility of low response rates at the time of phone surveys.

[4]This stratification based on the school vector chosen in distance bin 1, satisfies the Equal Treatment of Equals (ETE) property following Abdulkadiroğlu, Angrist, Narita and Pathak (2017). The ETE property is satisfied as all applicants who chose the exact same combination of schools in bin 1 are treated equally at the time of each school's randomization. They are subjected to the same randomization at each school which is listed in the vector, until they get matched at a school. Thus, on average the winners and non-winners who chose the same school vector in bin 1, are comparable to each other.

5. Finally, from each school vector which satisfies the above three criteria, I perform a stratified random sampling where the two strata are winners and non-winners corresponding to a given school vector chosen in bin 1. Furthermore, the sampling is done such that the count of applicants sampled per school vector = min(winners, non-winners, 25)* 2.[5]

[5] I restrict the maximum count of applicants per vector in order to maximize the count of unique school vectors in my sample. Based on all these criteria, the minimum number of applicants selected per school vector is equal to 8. Importantly, when the school vector consists of multiple schools chosen in bin 1, I make sure to sample a non-zero count of winning applicants (among winners) from each school in the vector.

A.3.3 Calculation of ex-ante propensity scores of winning under the lottery mechanism

Below, I explain the step-by-step process for calculating the simulated ex-ante propensity scores of winning under Maharashtra's lottery mechanism for RTE.

1. I conduct a large number of simulations of the lottery mechanism as explained in Section 1.3.1 (N $\sim$ 10,000).
2. For each simulation, I record the school allotted to each child.

Then for each child, I compute:

1. Simulated ex-ante probability of winning at *each* school that the child listed in application. I do this by averaging across simulations, the probability of winning at that school.
2. Simulated ex-ante probability of winning *in bin 1*. This is given by the sum of simulated ex-ante probability of winning at each school that the child listed in bin 1. The individual simulated probabilities for each chosen school are computed in the previous step.
3. Simulated ex-ante probability of winning *in any bin*. This is given by the sum of simulated ex-ante probability winning at each school that the child listed (combining bin 1, bin 2, bin 3).
4. Simulated ex-ante probability of winning at *elite schools*. I have two measures of elitness - PCA based index and school fee-based measure.

   i. For each child in the sample, and correspondingly for each RTE school that they listed in their application, I make an indicator of whether the school is elite or budget, based on the percentile cutoff. I code the indicator variable = 1 if the school lies above the respective percentile cutoff value, and I code it = 0 if the schools lies below the respective percentile cutoff value. The indicator variable is assiged a missing value in the case where there is missing data on PCA index or fee for the school.

ii. Next, I compute the simulated ex-ante propensity of winning at elite schools. To do this I simply take the sum of the simulated ex-ante propensities for each school that is coded to be elite based on the respective percentile cutoff.

5. Note that this is always satisfied: simulated probability $\in$ [0,1]

6. Next, I divide these into 100 bins of width = .01 each (for some estimations I reduce the number of bins to 50, in which case the width becomes .02, respectively).

7. Finally, I create dummies of narrow bins of simulated ex-ante propensity scores. In the case where I have 100 bins of propensity scores, this creates 100 dummies of narrow bins: [0,0.01] , (0.01,0.02], ..., (0.99, 1], such that only one of these 100 dummies gets activated for each applicant child.

8. In the estimations I control for dummies of narrow bins of ex-ante propensities of winning, as this facilitates the within-comparison between ex-ante similar applicants who vary in their lottery outcome.

A.4 Estimating Complier Characteristics and Counterfactual Destinies

A.4.1 Estimation

I follow the Angrist, Hull and Walters (2023)'s implementation of methods used in Abadie (2002) to compute complier characteristics and counterfactual destinies for untreated compliers. Below I discuss the steps as mentioned in Angrist, Hull and Walters (2023).

The notation is as follows: $Z_i \in \{0,1\}$ is the instrument which denotes whether i wins the RTE private school lottery. $D_i(1)$ and $D_i(0)$ refer to potential treatments, indicating i's RTE enrollment status as a quota student, when $Z_i = 1$ and $Z_i = 0$, respectively. $Y_i(0)$ and $Y_i(1)$ denote the potential outcomes for individual i as a function of RTE enrollment.

The following assumptions are made:

Assumption 1. Independence/exclusion: $(Y_i(0), Y_i(1), D_i(0), D_i(1)) \perp Z_i$.

Assumption 2. First stage: $\mathbb{E}[D_i|Z_i = 1] > \mathbb{E}[D_i|Z_i = 0]$.

Assumption 3. Monotonicity: $D_i(1) \geq D_i(0)$ $\forall$ i.

Angrist, Hull and Walters (2023) explain the process of backing out complier characteristics, which I discuss next. While individual compliers are not coded in any data, complier characteristics can be described using methods of Abadie (2002). The monotonicity assumption implies that the population contributing to the IV analysis only consists of always-takers, never-takers, and compliers. Some of the always and never takers can be identified by the following cells of the data: $D_i = 0$ and $Z_i = 1$ are always-takers while, $D_i = 1$ and $Z_i = 0$ are never-takers. The other cells of the data contain mixtures of compliers with the other two groups: $D_i = 0$ and $Z_i = 0$ contain compliers and never-takers, while $D_i = 1$ and $Z_i = 1$ contain compliers and always-takers. The size of the compliers is given by the first stage. The data also helps in infering the share of never-takers and always-takers as these correspond to the proportion of those who reject the offer of enrollment as a quota student, and the proportion of those who choose to enroll as a quota student when not offered.

Like them, I estimate the following system of equations via 2SLS

$$g(X_i, Y_i) \times 1\{D_i = d\} = \pi_d + \gamma_d 1\{D_i = d\} + v_{id} \tag{A.1}$$

$$1\{D_i = d\} = \phi_d + \beta_d Z_i + e_{id}, d \in \{0, 1\} \tag{A.2}$$

, where $g(X_i, Y_i)$ is a function of student baseline characteristics (X_i) or post-lottery outcomes (Y_i). Complier characteristics for the treated are obtained by setting $d = 1$ which amounts to using Z_i as the instrument for D_i where the outcome in the second stage is given by $g(X_i, Y_i)$ multiplied by D_i. Similarly setting $d = 0$, estimates the complier characteristics for the untreated which means using Z_i as an instrument for (1-D_i) where the outcome in the second stage is $g(X_i, Y_i)$ multiplied by (1-D_i).

Estimating complier characteristics: Setting $g(X_i, Y_i) = X_i$ yeilds the average complier characteristics for baseline covariates. Estimating equations (A.1) and (A.2) as explained in the previous paragraph (along with ex-ante propensities of winning) produces the columns (1) and (2) for Table A.4. Column (3) shows always-taker means which are computed by regressing $X_i D_i (1 - Z_i)$ on $D_i (1 - Z_i)$ (with ex-ante propensities), column (4) shows never-taker means which are computed by regressing $X_i D_i Z_i$ on $(1 - D_i) Z_i$ (with ex-ante propensities).

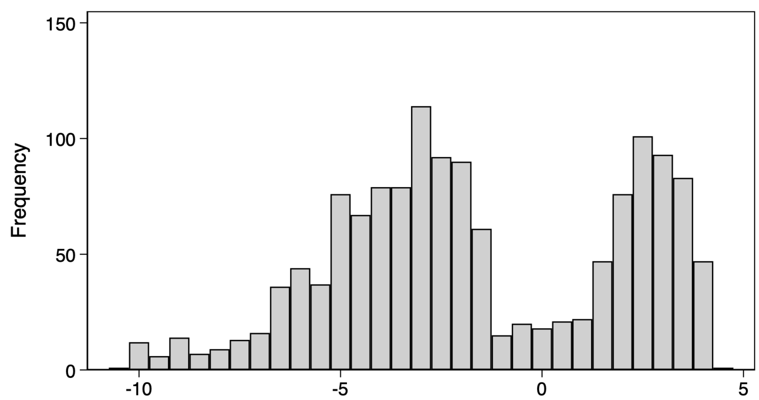
Estimating counterfactual destinies: Table A.5 shows the distribution of enrollment across sectors for lottery losers. Lottery losers could be enrolled at private schools as fee-paying students, government schools, or remain out-of-school. I first create dummies of enrollment at a particular school sector. Next, I estimate (A.1) and (A.2) by setting d=0, for a total of four times (since there are four outside options), each time setting $g(X_i, Y_i)$ as the dummy for enrollment at that specific outisde option.

Appendix B

Appendix: Chapter 2

B.1 Figures

Figure B.1: Distribution of age difference between applicant and their sibling.



Notes: This is a histogram showing the distribution of age difference between applicant child and their sibling.

B.2 Tables

Table B.1: Impact of being in a winning household on sibling's enrollment

	Enrollment (2020-21)		Enrollment (2021-22)	
	Any	Private	Any	Private
	(1)	(2)	(3)	(4)
WinningHH	-0.027	-0.005	-0.004	0.015
	(0.018)	(0.027)	(0.017)	(0.025)
Dependent mean	0.84	0.84	0.83	0.83
Observations	1,198	1,172	1,358	1,351
R^2	0.42	0.20	0.44	0.21
Pscores of winning	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Sample comprises siblings of RTE 25% quota applicants. Age categories refer to sibling's age at the time of survey (in Nov, 2021) and siblings are between 4 years and 18 years old at the time of survey. Retrospective information on enrollment is collected for two academic years for each sibling. The sample size is smaller in column 1 because the enrollment status of siblings is asked only of those who are 4-18 years in the corresponding school year. Hence, 4 year olds in 2021 have missing information on enrollment in the 2020-21 school year. Column 2 includes all surveyed siblings. The reference age category comprises children who are older than applicant child and younger than or equal to 10 years old in 2021.

Table B.2: Impact of being in a winning household on sibling's enrollment: treatment effect heterogeneity by sibling's age

	Enrollment (2020-21)		Enrollment (2021-22)	
	Any	Private	Any	Private
	(1)	(2)	(3)	(4)
WinningHH	-0.010	0.025	-0.008	0.062
	(0.034)	(0.053)	(0.035)	(0.053)
Age=[4]			-0.633***	-0.515***
			(0.039)	(0.060)
Age=[5]	-0.495***	-0.414***	-0.421***	-0.311***
	(0.038)	(0.060)	(0.038)	(0.059)
Age=[6,applicantage)	-0.384***	-0.405***	-0.340***	-0.303***
	(0.043)	(0.069)	(0.043)	(0.066)
Age=[11,12]	0.019	-0.054	-0.006	-0.053
	(0.031)	(0.048)	(0.031)	(0.048)
Age=[13,18]	-0.003	-0.090*	-0.023	-0.090*
	(0.034)	(0.053)	(0.034)	(0.052)
WinningHH*Age=[4]			-0.032	-0.088
			(0.058)	(0.089)
WinningHH*Age=[5]	-0.153***	-0.131	-0.060	-0.080
	(0.056)	(0.089)	(0.057)	(0.087)
WinningHH*Age=[6,applicantage)	0.013	0.078	0.093	0.018
	(0.066)	(0.105)	(0.066)	(0.101)
WinningHH*Age=[11,12]	0.009	-0.031	0.011	-0.066
	(0.048)	(0.075)	(0.048)	(0.074)
WinningHH*Age=[13,18]	0.014	-0.033	0.028	-0.055
	(0.051)	(0.079)	(0.052)	(0.079)
Dependent mean	0.84	0.64	0.83	0.83
Observations	1,198	1,172	1,358	1,351
R^2	0.43	0.20	0.44	0.21
Pscores of winning	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Sample comprises siblings of RTE 25% quota applicants. Age categories refer to sibling's age at the time of survey (in Nov, 2021) and siblings are between 4 years and 18 years old at the time of survey. Retrospective information on enrollment is collected for two academic years for each sibling. The sample size is smaller in the first two columns because enrollment status is asked only of those who are 4-18 years old in the corresponding school year. Hence, 4 year olds in 2021 have missing information on enrollment in the 2020-21 school year. Column (3) and (4) include all surveyed siblings. The reference age category comprises children who are older than the applicant child and younger than or equal to 10 years old in 2021.

Table B.3: Coefficients of interest (from Table B.2)

		Enrollment (2020-21)		Enrollment (2021-22)	
		Any	Pvt	Any	Pvt
Age (in 2021)	Coefficient	(1)	(2)	(3)	(4)
Age 4	$\alpha_2 + \alpha_8$			-0.039	-0.025
				(0.047)	(0.072)
Age 5	$\alpha_2 + \alpha_9$	-0.162***	-0.106	-0.067	-0.018
		(0.044)	(0.071)	(0.045)	(0.069)
$6 \leq$ Age $\leq$ applicant	$\alpha_2 + \alpha_{10}$	0.003	0.103	0.085	0.079
		(0.056)	(0.090)	(0.056)	(0.085)

Notes: This table presents the coefficients of interest from Table B.2.

Table B.4: Impact of being in a winning household on sibling's access to educational resources

	School provides instruction (2020-21) (own school)	School provides instruction (2020-21) (own or RTE applicant's school)
	(1)	(2)
WinningHH	0.036	0.043
	(0.039)	(0.035)
Age=[4]		
Age=[5]	-0.564***	-0.196***
	(0.043)	(0.041)
Age=[6,applicantage)	-0.505***	-0.243***
	(0.049)	(0.045)
Age=[11,12]	0.033	0.031
	(0.035)	(0.032)
Age=[13,18]	0.039	0.051
	(0.039)	(0.035)
WinningHH*Age=[4]		
WinningHH*Age=[5]	-0.131**	0.069
	(0.064)	(0.059)
WinningHH*Age=[6,applicantage)	-0.041	0.026
	(0.075)	(0.068)
WinningHH*Age=[11,12]	0.007	-0.017
	(0.055)	(0.049)
WinningHH*Age=[13,18]	-0.017	-0.035
	(0.058)	(0.052)
Dependent mean	0.84	0.87
Observations	1,198	1,159
R^2	0.45	0.19
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: The outcome in column (1) is an indicator for whether the sibling's school was providing any instruction in the 2020-21 school year; it is set equal to zero for those not enrolled in any school. The outcome in column (2) indicates whether the sibling's own school or the RTE applicant child's school provides any instruction in the 2020-21 school-year. Sample comprises siblings of RTE 25% quota applicants. Age categories refer to sibling's age at the time of survey (in Nov, 2021) and siblings are between 4 years and 18 years old at the time of survey. Hence, those who are 4 years old in 2021 are excluded here as they have missing information on enrollment in the 2020-21 school year. The reference age category comprises children who are older than the applicant child and younger than or equal to 10 years old in 2021.

Table B.5: Coefficients of interest (from Table B.4)

		School provides instruction (2020-21) (own school)	School provides instruction (2020-21) (own or RTE applicant's school)
Age (in 2021)	Coefficient	(1)	(2)
Age 5	$\alpha_2 + \alpha_9$	-0.095*	0.112**
		(0.051)	(0.047)
$6 \leq$ Age $\leq$ applicant	$\alpha_2 + \alpha_{10}$	-0.004	0.068
		(0.063)	(0.058)

Notes: This table presents the coefficients of interest from Table B.4.

Table B.6: Impact on parental monetary investments in sibling

	Any expense on		
	School fees (Yes/No)	Tutoring Fee (Yes/No)	Other educational needs (Yes/No)
	(1)	(2)	(3)
WinningHH	0.015	-0.019	0.008
	(0.027)	(0.027)	(0.020)
Dependent mean	0.63	0.40	0.85
Dependent mean (control)	0.62	0.42	0.85
Observations	1,351	1,358	1,319
R^2	0.08	0.09	0.08
Pscores of winning	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: The outcomes measure the extensive margin of any educational expenses on the child in the past one year (on school fee, after school private tutoring, curriculum books, notebooks and stationary). Sample comprises siblings of RTE 25% quota applicants. Age categories refer to sibling's age at the time of survey (in Nov, 2021) and siblings are between 4 years and 18 years at the time of survey.

Table B.7: Impact on parental monetary investments in sibling by age groups

	Any expense on		
	School fees (Yes/No)	Tutoring Fee (Yes/No)	Other educational needs (Yes/No)
	(1)	(2)	(3)
WinningHH	0.062	0.007	-0.001
	(0.053)	(0.057)	(0.040)
Age=[4]	-0.515***	-0.201***	-0.359***
	(0.060)	(0.064)	(0.045)
Age=[5]	-0.311***	-0.081	-0.155***
	(0.059)	(0.063)	(0.044)
Age=[6,applicantage)	-0.303***	-0.096	-0.136***
	(0.066)	(0.071)	(0.049)
Age=[11,12]	-0.053	-0.031	-0.011
	(0.048)	(0.051)	(0.036)
Age=[13,18]	-0.090*	-0.064	-0.003
	(0.052)	(0.056)	(0.040)
WinningHH*Age=[4]	-0.088	-0.035	0.043
	(0.089)	(0.096)	(0.067)
WinningHH*Age=[5]	-0.080	-0.029	-0.040
	(0.087)	(0.093)	(0.066)
WinningHH*Age=[6,applicantage)	0.018	0.026	0.059
	(0.101)	(0.108)	(0.076)
WinningHH*Age=[11,12]	-0.066	-0.088	0.017
	(0.074)	(0.079)	(0.056)
WinningHH*Age=[13,18]	-0.055	0.005	0.016
	(0.079)	(0.085)	(0.060)
Dependent mean	0.63	0.40	0.85
Dependent mean (control)	0.62	0.42	0.85
Observations	1,351	1,358	1,319
R^2	0.21	0.11	0.18
Pscores of winning	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: The outcomes measure the extensive margin of any educational expenses on the child in the past one year (on school fee, after school private tutoring, curriculum books, notebooks and stationary). Sample comprises siblings of applicants and the age categories refer to sibling's age at the time of survey (in 2021). The reference age category comprises children who are older than the applicant child and younger than or equal to 10 years old in 2021.

Table B.8: Impact on parental time investments in sibling

	Any help (Yes/No)	Hours of help (hrs/week)
	(1)	(2)
WinningHH	-0.019	-0.219
	(0.027)	(0.328)
Dependent mean	0.62	0.62
Dependent mean (control)	0.63	0.63
Observations	1,358	1,358
R^2	0.10	0.08
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: Column (1) measures the extensive margin of whether the child receives any help with educational activities in the household, and column (2) measures the intensive margin of the number of hours of help. The survey questions were: "Does the child receive any help with educational activities from any members of the household?" followed by details of each person who helps and their relationship with the child. Next, it was asked: "Among all those who help, who is the person who most often helps the child with educational activities?", followed by details about the number of hours of help on a typical day, and the number of days of help per week in the past week, to calculate weekly hours of help coming from the main helper. Sample comprises siblings of RTE 25% quota applicants. Age categories refer to sibling's age at the time of survey (in Nov, 2021) and siblings are between 4 years and 18 years old at the time of survey.

Table B.9: Impact on parental time investments in sibling by age groups

	Any help (Yes/No)	Hours of help (hrs/week)
	(1)	(2)
WinningHH	-0.063	-0.173
	(0.050)	(0.656)
Age=[4]	0.193***	0.650
	(0.057)	(0.736)
Age=[5]	0.258***	2.353***
	(0.055)	(0.718)
Age=[6,applicantage)	0.224***	1.465*
	(0.062)	(0.813)
Age=[11,12]	-0.120***	-1.120*
	(0.045)	(0.588)
Age=[13,18]	-0.421***	-4.228***
	(0.049)	(0.643)
WinningHH*Age=[4]	0.075	-0.249
	(0.084)	(1.102)
WinningHH*Age=[5]	0.006	-0.663
	(0.082)	(1.072)
WinningHH*Age=[6,applicantage)	0.051	-0.105
	(0.096)	(1.247)
WinningHH*Age=[11,12]	-0.023	-0.774
	(0.070)	(0.915)
WinningHH*Age=[13,18]	0.172**	1.301
	(0.075)	(0.975)
Dependent mean	0.62	5.74
Dependent mean (control)	0.63	5.82
Observations	1,358	1,358
R^2	0.29	0.17
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: Column (1) measures the extensive margin of whether the child receives any help with educational activities in the household, and column (2) measures the intensive margin of the number of hours of help. The survey questions were: "Does the child receive any help with educational activities from any members of the household?" followed by details of each person who helps and their relationship with the child. Next, it was asked: "Among all those who help, who is the person who most often helps the child with educational activities?", followed by details about the number of hours of help on a typical day, and the number of days of help per week in the past week, to calculate weekly hours of help coming from the main helper. Sample comprises siblings of RTE 25% quota applicants. Age categories refer to sibling's age at the time of survey (in Nov, 2021) and siblings are between 4 years and 18 years old at the time of survey. The reference age category comprises children who are older than the applicant child and younger than or equal to 10 years old in 2021.

Table B.10: Treatment effect heterogeneity in enrollment by quota type

	Enrollment (2020-21)	Enrollment (2021-22)
	(1)	(2)
WinningHH	-0.041	0.001
	(0.026)	(0.025)
Low Income Quota	-0.088	0.013
	(0.061)	(0.061)
WinningHH*Low Income Quota	0.026	-0.030
	(0.048)	(0.046)
Dependent mean	0.84	0.83
Observations	1,198	1,358
R^2	0.09	0.08
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: This table looks at treatment effect heterogeneity in enrollment by whether the household applied to the RTE policy under the low income quota or disadvantaged caste quota.

Table B.11: Treatment effect heterogeneity in enrollment by quota type and age

	Enrollment 2020-21	Enrollment 2021-22
	(1)	(2)
WinningHH	-0.015	-0.003
	(0.042)	(0.042)
Low Income Quota	0.045	0.009
	(0.064)	(0.064)
Age=[4]		-0.685***
		(0.047)
Age=[5]	-0.464***	-0.421***
	(0.045)	(0.046)
Age=[6,applicantage)	-0.343***	-0.369***
	(0.053)	(0.053)
Age=(11,12]	0.040	-0.009
	(0.037)	(0.038)
Age=(13,18]	0.023	-0.013
	(0.041)	(0.041)
WinningHH*Low Income Quota	0.029	-0.017
	(0.074)	(0.074)
WinningHH*Age=[4]		0.039
		(0.069)
WinningHH*Age=[5]	-0.116*	-0.037
	(0.068)	(0.068)
WinningHH*Age=[6,applicantage)	0.016	0.107
	(0.079)	(0.080)
WinningHH*Age=(11,12]	0.005	0.017
	(0.059)	(0.059)
WinningHH*Age=(13,18]	0.005	-0.001
	(0.061)	(0.062)
Low Income Quota*Age=[4]		0.206**
		(0.087)
Low Income Quota*Age=[5]	-0.103	-0.002
	(0.084)	(0.086)
Low Income Quota*Age=[6,applicantage)	-0.118	0.081
	(0.090)	(0.090)
Low Income Quota*Age=(11,12]	-0.061	0.010
	(0.067)	(0.068)
Low Income Quota*Age=(13,18]	-0.076	-0.040
	(0.075)	(0.075)
WinningHH*Age=[4]*Low Income Quota		-0.280**
		(0.132)
WinningHH*Age=[5]*Low Income Quota	-0.120	-0.065
	(0.123)	(0.124)
WinningHH*Age=[6,applicantage)*Low Income Quota	-0.033	-0.036
	(0.141)	(0.142)
WinningHH*Age=(11,12]*Low Income Quota	-0.003	-0.014
	(0.104)	(0.105)
WinningHH*Age=(13,18]*Low Income Quota	0.028	0.113
	(0.114)	(0.115)
Dependent mean	0.84	0.83
Observations	1,198	1,358
R^2	0.43	0.45
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: This table looks at treatment effect heterogeneity in enrollment by sibling's age and by whether the household applied to the RTE policy under the low income quota or disadvantaged caste quota.

Table B.12: Coefficients of interest (from Table B.11)

Age (in 2021)	Coefficient	Enrollment 2020-21 (1)	Enrollment 2021-22 (2)
Panel A: For disadvantaged caste households			
Age 4	$\beta_2 + \beta_{10}$		0.042
			(0.054)
Age 5	$\beta_2 + \beta_{11}$	-0.128**	-0.038
		(0.053)	(0.053)
$6 \leq$ Age $\leq$ applicant	$\beta_2 + \beta_{12}$	0.006	0.104
		(0.067)	(0.067)
		(1)	(2)
Panel B: For low income households			
Age 4	$\beta_2 + \beta_9 + \beta_{10} + \beta_{20}$		-0.255***
			(0.094)
Age 5	$\beta_2 + \beta_9 + \beta_{11} + \beta_{21}$	-0.217***	-0.116
		(0.082)	(0.083)
$6 \leq$ Age $\leq$ applicant	$\beta_2 + \beta_9 + \beta_{12} + \beta_{22}$	-0.005	0.051
		(0.101)	(0.101)

Notes: This table presents the coefficients of interest from Table B.11.

Table B.13: Treatment effect heterogeneity in enrollment by sibling's gender

	Enrollment (2020-21)	Enrollment (2021-22)
	(1)	(2)
WinningHH	-0.064*	-0.016
	(0.034)	(0.032)
Male	0.026	0.009
	(0.029)	(0.028)
WinningHH*Male	0.056	0.014
	(0.045)	(0.043)
Dependent mean	0.84	0.83
Observations	1,198	1,358
R^2	0.08	0.07
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: This table looks at treatment effect heterogeneity in enrollment by sibling's gender.

Table B.14: Treatment effect heterogeneity in enrollment by sibling's gender and age

	Enrollment 2020-21	Enrollment 2021-22
	(1)	(2)
WinningHH	-0.011	0.005
	(0.053)	(0.054)
Male	-0.005	0.016
	(0.045)	(0.046)
Age=[4]		-0.577***
		(0.058)
Age=[5]	-0.539***	-0.393***
	(0.055)	(0.057)
Age=[6,applicantage)	-0.357***	-0.334***
	(0.060)	(0.061)
Age=(11,12]	0.013	0.007
	(0.046)	(0.047)
Age=(13,18]	-0.003	-0.026
	(0.051)	(0.052)
WinningHH*Male	0.014	-0.017
	(0.069)	(0.071)
WinningHH*Age=[4]		-0.021
		(0.087)
WinningHH*Age=[5]	-0.096	-0.059
	(0.083)	(0.085)
WinningHH*Age=[6,applicantage)	-0.336***	0.005
	(0.103)	(0.105)
WinningHH*Age=(11,12]	0.002	-0.010
	(0.073)	(0.075)
WinningHH*Age=(13,18]	-0.012	0.006
	(0.078)	(0.080)
Male*Age=[4]		-0.093
		(0.078)
Male*Age=[5]	0.098	-0.035
	(0.075)	(0.076)
Male*Age=[6,applicantage)	-0.043	0.010
	(0.085)	(0.087)
Male*Age=(11,12]	0.017	-0.022
	(0.061)	(0.063)
Male*Age=(13,18]	0.014	0.008
	(0.068)	(0.069)
WinningHH*Age=[4]*Male		-0.019
		(0.118)
WinningHH*Age=[5]*Male	-0.118	-0.006
	(0.112)	(0.114)
WinningHH*Age=[6,applicantage)*Male	0.529***	0.123
	(0.134)	(0.137)
WinningHH*Age=(11,12]*Male	0.004	0.037
	(0.096)	(0.098)
WinningHH*Age=(13,18]*Male	0.045	0.035
	(0.103)	(0.105)
Dependent mean	0.84	0.83
Observations	1,198	1,358
R^2	0.44	0.44
Pscores of winning	Yes	Yes
Controls	Yes	Yes

Notes: This table looks at treatment effect heterogeneity in enrollment by sibling's gender. The reference age category comprises children who are older than applicant child and younger than 10 years in 2021-22.

Table B.15: Coefficients of interest (from Table B.14)

Age (in 2021)	Coefficient	Enrollment 2020-21 (1)	Enrollment 2021-22 (2)
Panel A: For female siblings			
Age 4	$\beta_2 + \beta_{10}$		-0.015
			(0.068)
Age 5	$\beta_2 + \beta_{11}$	-0.107*	-0.053
		(0.064)	(0.065)
$6 \leq$ Age $\leq$ applicant	$\beta_2 + \beta_{12}$	-0.346***	0.010
		(0.088)	(0.089)
		(1)	(2)
Panel B: For male siblings			
Age 4	$\beta_2 + \beta_9 + \beta_{10} + \beta_{20}$		-0.051
			(0.064)
Age 5	$\beta_2 + \beta_9 + \beta_{11} + \beta_{21}$	-0.210***	-0.076
		(0.060)	(0.062)
$6 \leq$ Age $\leq$ applicant	$\beta_2 + \beta_9 + \beta_{12} + \beta_{22}$	0.196***	0.115
		(0.073)	(0.074)

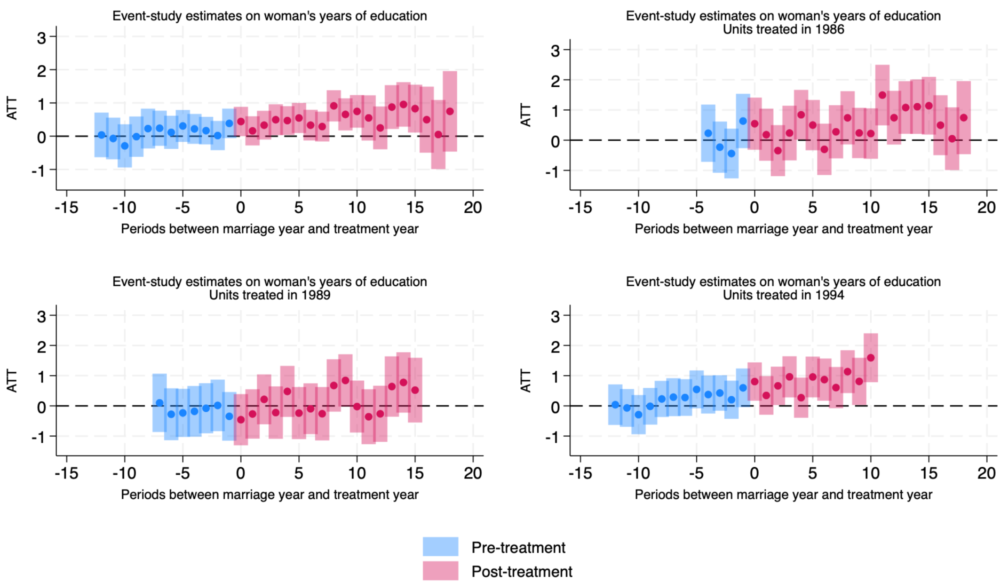
Notes: This table presents the coefficients of interest from Table B.14.

Appendix C

Appendix: Chapter 3

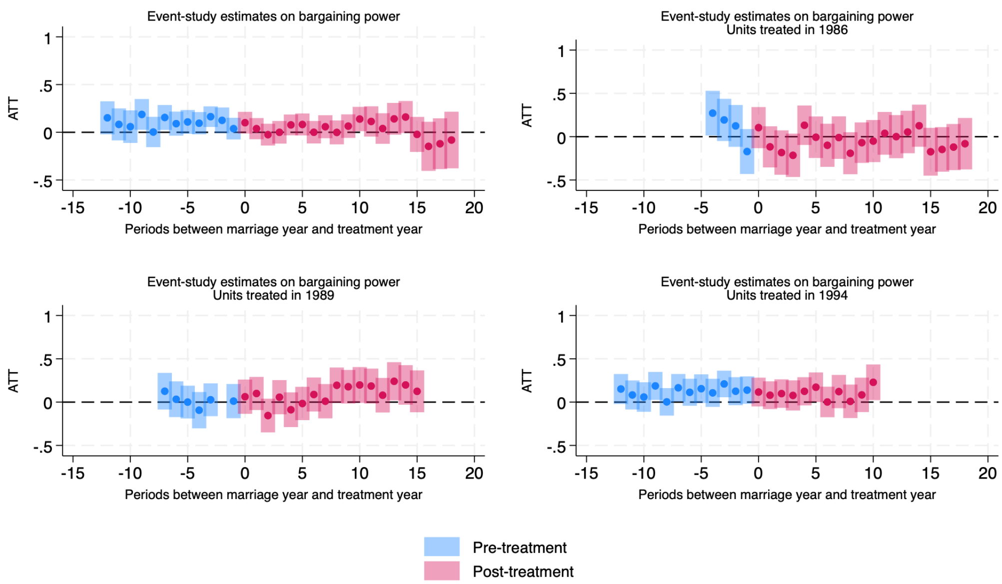
C.1 Figures

Figure C.1: Event study on years of Education (Wife)



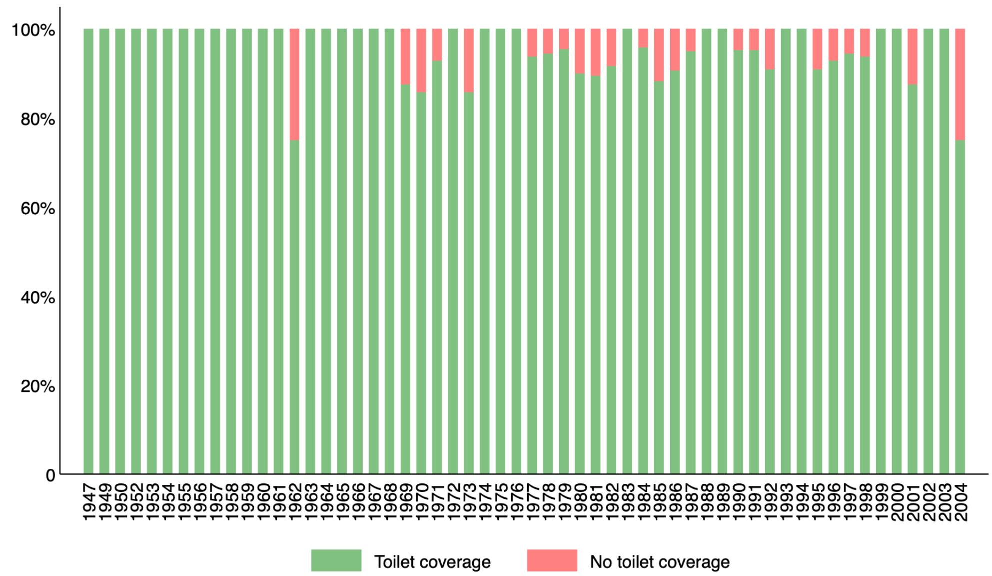
Notes: This figure plots event study estimates of the impact of policy on women's years of educational attainment, by comparing marriage cohorts over time and using never-treated states as comparison groups.

Figure C.2: Event study on bargaining power (Wife)



Notes: This figure plots event study estimates of the impact of policy on women's intra-household bargaining power, by comparing marriage cohorts over time and using never-treated states as comparison groups.

Figure C.3: Toilet prevalence by year of marriage in Kerala



Notes: Green bars show the unconditional likelihood of the presence of a toilet in the household against the year of marriages that happened in Kerala. The corresponding red bars correspond to the unconditional likelihood of the absence of a toilet. The year 1975 is the year before the HSA amendment was implemented in Kerala. We see that in 1975, the year before the HSAA was implemented, 100% of the households had a toilet. Note that our data can only tell us whether the household has a toilet in 2005 (the survey year), and it does not give any information about the year in which the toilet was constructed. So the plots here represents whether these households had a toilet in 2005.

C.2 Tables

Table C.1: Summary Statistics

	All states	Reform states	Non-Reform states
Panel A: Sample of Wife of the household head			
Toilet	0.34	0.35	0.33
	(0.47)	(0.47)	(0.47)
Years of education of head	7.01	7.20	6.93
	(5.24)	(5.19)	(5.26)
Years of education of wife	4.92	5.71	4.58
	(5.18)	(5.11)	(5.18)
Bargaining power in marital household	-0.02	0.01	-0.03
	(1.02)	(1.00)	(1.02)
Age at marriage	15.49	16.13	15.34
	(5.6)	(5.41)	(5.06)
Urban	0.46	0.47	0.45
	(0.49)	(0.48)	(0.48)
Wealth index	-0.01	0.07	-0.04
	(1.01)	(0.93)	(1.04)
Scheduled caste	0.22	0.19	0.23
	(0.41)	(0.39)	(0.42)
Scheduled tribe	0.08	0.06	0.09
	(0.28)	(0.24)	(0.29)
Other backward class (OBC)	0.36	0.50	0.29
	(0.48)	(0.50)	(0.45)
General caste	0.31	0.22	0.35
	(0.46)	(0.41)	(0.47)
Observations	32,169	9,932	22,237
Panel B: Sample of Any married woman			
Toilet	0.36	0.36	0.36
	(0.48)	(0.48)	(0.48)
Years of education of head	6.35	6.58	6.26
	(5.29)	(5.25)	(5.31)
Years of education of woman	5.32	5.90	5.09
	(5.28)	(5.14)	(5.32)
Bargaining power in marital household	-0.04	0.01	-0.06
	(1.01)	(0.99)	(1.02)
Age at marriage	18.66	18.72	18.64
	(3.66)	(3.81)	(3.59)
Urban	0.44	0.46	0.43
	(0.49)	(0.48)	(0.47)
Wealth index	-0.01	0.07	-0.04
	(1.02)	(0.94)	(1.05)
Scheduled caste	0.22	0.19	0.23
	(0.41)	(0.39)	(0.42)
Scheduled tribe	0.08	0.07	0.09
	(0.28)	(0.25)	(0.28)
Other backward class (OBC)	0.36	0.49	0.31
	(0.48)	(0.49)	(0.46)
General caste	0.33	0.23	0.37
	(0.47)	(0.42)	(0.48)
Observations	42,765	12,689	30,076

Notes: In our estimation, we control for the following pre-treatment characteristics: indicator for urban, caste categories, and household wealth index.

Table C.2: Impact of HSAA on women's years of educational attainment (Wife)

	(1) Never treated	(2) Not yet treated
Aggregate ATT (GAverage)	0.558^{***}	0.559^{***}
	(0.197)	(0.197)
ATT of units treated in 1986	0.527	0.507
	(0.376)	(0.370)
ATT of units treated in 1989	0.105	0.129
	(0.368)	(0.363)
ATT of units treated in 1994	0.794^{***}	0.794^{***}
	(0.290)	(0.290)
Observations	32,169	32,169

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). We present estimates using never-treated as comparison. Standard errors are computed using wild cluster bootstrap at the state level. The last panel of the table reports estimates of a chi-square test which tests the null hypothesis of no differential pre-trends between treated and untreated units.

Table C.3: Impact of HSAA on women's intra-household bargaining power (Wife)

	(1) Joint	(2) Decision-making	(3) Mobility	(4) Financial	(5) Low IPV
Aggregate ATT (GAverage)	0.0602	0.102^{**}	0.0570	-0.00590	-0.0440
	(0.0513)	(0.0515)	(0.0494)	(0.0502)	(0.0528)
ATT of units treated in 1986	-0.0500	0.0246	-0.0644	-0.108	-0.0356
	(0.109)	(0.110)	(0.103)	(0.0907)	(0.115)
ATT of units treated in 1989	0.0873	0.169^{*}	0.107	-0.0693	-0.0453
	(0.0895)	(0.0973)	(0.0804)	(0.0934)	(0.107)
ATT of units treated in 1994	0.100	0.107	0.0917	0.0756	-0.0474
	(0.0739)	(0.0713)	(0.0732)	(0.0759)	(0.0709)
Observations	32,169	32,169	32,169	32,169	32,169

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). We present estimates using never-treated as comparison. Standard errors are computed using wild cluster bootstrap at the state level.

Table C.4: Impact of HSAA on toilet ownership (Rural sample only)

	(1) Never treated	(2) Not yet treated
Aggregate ATT (GAverage)	0.0300	0.0306
	(0.0202)	(0.0203)
ATT of units treated in 1986	0.0262	0.0270
	(0.0461)	(0.0457)
ATT of units treated in 1989	-0.00108	0.000505
	(0.0359)	(0.0354)
ATT of units treated in 1994	0.0468^{*}	0.0468^{*}
	(0.0284)	(0.0284)
Observations	19,122	19,122

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the coefficient estimates and standard errors of each treated group's average treatment effect on the treated parameter following Callaway and Sant'Anna (2021). Standard errors are computed using wild cluster bootstrap at the state level.

Table C.5: Impact of HSAA on toilet ownership (Two way fixed effects)

	Toilet ownership
	(1)
Treated	0.022***
	(0.009)
Observations	32,169
R^2	0.45
State FE	Yes
Year of marriage FE	Yes
Controls	Yes

Notes: This table reports the coefficient estimates and standard errors using two-way fixed effects.

C.3 Proof of Proposition 1

In this subsection we show formal proof of Proposition 1 in the text.

Proof. We start by re-iterating that over some set of comparison groups $\mathcal{G}_{\text{comp}}$ such that $g' > t$ for all $g' \in \mathcal{G}_{\text{comp}}$, the above assumptions identify the true group-time treatment effects if both the group identity G_i and the treatment eligibility b_i are observed. In this case the true $ATT(g,t)$ is given by

$$ATT(g,t) = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}, b_i = 1\right]$$

However, since we do not observe b_i for all units i, we can identify (and estimate) the following expression, which we denote as $ATT^*(g,t)$

$$ATT^*(g,t) = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\right]$$

Now using the Law of Iterated Expectations, we rewrite the above identified expression as,

$$\begin{aligned}ATT^*(g,t) = {} & \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1\right] \mathbb{P}(b_i = 1 \mid G_i = g) \\ & - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}, b_i = 1\right] \mathbb{P}(b_i = 1 \mid G_i \in \mathcal{G}_{\text{comp}})\end{aligned}$$

By our assumption that the event b_i is independent of group indicators, we have

$$\begin{aligned}ATT^*(g,t) &= \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1\right] \mathbb{P}(b_i = 1) - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}, b_i = 1\right] \mathbb{P}(b_i = 1) \\ &= \mathbb{P}(b_i = 1)\left(\mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}, b_i = 1\right]\right) \\ &= \mathbb{P}(b_i = 1) ATT(g,t)\end{aligned}$$

since $\mathbb{P}(b_i = 1) \in [0,1]$, we have that

$$\mid ATT^*(g,t) \mid \leq \mid ATT(g,t) \mid$$

Hence, if the true treatment effect $ATT(g,t)$ is positive then $ATT^*(g,t) \leq ATT(g,t)$.

This proof can be easily extended to a case where we also condition on other covariates

X_i which are independent of b_i and G_i. In this case, under the assumption of conditional parallel trends based on comparison group $\mathcal{G}_{comp}$, along with the assumptions on random sampling and no anticipation, we can write the true $ATT(g,t)$ as

$$ATT(g,t) = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1, X_i\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\ , b_i = 1, X_i\right]$$

and the identified $ATT^*(g,t)$ given the data limitation as

$$ATT^*(g,t) = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, X_i\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\ , X_i\right]$$

Using the Law of Iterated Expectations, we can write the above identified expression as,

$$\begin{aligned} ATT^*(g,t) = {} & \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1, X_i\right] \mathbb{P}(b_i = 1 \mid G_i = g, X_i) \\ & - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\ , b_i = 1, X_i\right] \mathbb{P}(b_i = 1 \mid G_i \in \mathcal{G}_{\text{comp}}\ , X_i) \end{aligned}$$

By our assumption that the event b_i is independent of other covariates and group indicators, we have

$$\begin{aligned} & ATT^*(g,t) \\ & = \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1, X_i\right] \mathbb{P}(b_i = 1 \mid X_i) - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\ , b_i = 1, X_i\right] \mathbb{P}(b_i = 1 \mid X_i) \\ & = \mathbb{P}(b_i = 1 \mid X_i) \left(\mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i = g, b_i = 1, X_i\right] - \mathbb{E}\left[Y_{i,t} - Y_{i,g-1} \mid G_i \in \mathcal{G}_{\text{comp}}\ , b_i = 1, X_i\right]\right) \\ & = \mathbb{P}(b_i = 1 \mid X_i) ATT(g,t) \\ & \leq ATT(g,t) \end{aligned}$$

Since $\mathbb{P}(b_i = 1 \mid X_i) \in [0,1]$, we have that

$$\mid ATT^*(g,t) \mid \ \leq \ \mid ATT(g,t) \mid$$

Hence, if the true treatment effect $ATT(g,t)$ is positive then $ATT^*(g,t) \leq ATT(g,t)$

Now let $\widehat{ATT(g,t)}$ be a consistent estimator of the true treatment effect $ATT(g,t)$. Hence if $ATT(g,t) \sim \mathcal{N}\left(\mu_g, \sigma_g^2\right)$, we have $\sqrt{n}\left(\widehat{ATT(g,t)} - \mu_g\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_g^2\right)$.

Now let $\widehat{p_x}$ be a consistent estimator of $\mathbb{P}(b_i = 1 \mid X_i)$. Using the Delta method, we have

$$\sqrt{n}\left(\widehat{p_x}\widehat{ATT(g,t)}\right) \xrightarrow{d} \mathcal{N}\left(\mathbb{P}(b_i = 1 \mid X_i)\mu_g, (\mathbb{P}(b_i = 1 \mid X_i)\sigma_g)^2\right)$$

Using the continuous mapping theorem, $\widehat{p_x}\widehat{ATT(g,t)}$ is a consistent estimator of $ATT^*(g,t)$. Thus,

$$ATT^*(g,t) \sim \mathcal{N}\left(\mathbb{P}(b_i = 1 \mid X_i)\mu_g, (\mathbb{P}(b_i = 1 \mid X_i)\sigma_g)^2\right)$$

It is straightforward to derive the asymptotic distribution of the average treatment effect $\hat{\beta}_1$ which is the parameter of interest.

$$\begin{aligned} &ATT(g,t) \sim \mathcal{N}\left(\mu_g, \sigma_g^2\right) \\ &\Rightarrow \sqrt{n}\left(\widehat{ATT(g,t)} - \mu_g\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma_g^2\right) \end{aligned}$$

Using the delta method, and that $ATT^*(g,t) = \mathbb{P}(b_i = 1 \mid X_i)ATT(g,t)$ we have

$$\sqrt{n}\left(\frac{\widehat{ATT(g,t)}}{\Pr(b_i = 1 \mid X_i)} - \frac{\mu_g}{\Pr(b_i = 1 \mid X_i)}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{\Pr(b_i = 1 \mid X_i)}\right)$$

Observe that the function $g(y) = \frac{y}{Pr(p=1|X)}$ is continuous and differentiable $\forall y \in \mathcal{R}$.

Hence, the estimated standard error is asymptotically an upper bound. Intuitively, this arises from the fact that the variance of the unobserved eligibility criterion remains as residual variance, thus reducing the precision of the estimator.

$\square$

Bibliography

Abadie, Alberto. 2002. "Bootstrap tests for distributional treatment effects in instrumental variable models." *Journal of the American statistical Association* 97(457):284–292.

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics* 113(2):231–263.

Abdulkadiroğlu, Atila, Joshua Angrist and Parag Pathak. 2014. "The elite illusion: Achievement effects at Boston and New York exam schools." *Econometrica* 82(1):137–196.

Abdulkadiroğlu, Atila, Joshua D Angrist, Susan M Dynarski, Thomas J Kane and Parag A Pathak. 2011. "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots." *The Quarterly Journal of Economics* 126(2):699–748.

Abdulkadiroğlu, Atila, Joshua D Angrist, Yusuke Narita and Parag A Pathak. 2017. "Research design meets market design: Using centralized assignment for impact evaluation." *Econometrica* 85(5):1373–1432.

Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg and Christopher R Walters. 2020. "Do parents value school effectiveness?" *American Economic Review* 110(5):1502–1539.

Adhvaryu, Achyuta and Anant Nyshadham. 2016. "Endowments at birth and parents' investments in children." *The Economic Journal* 126(593):781–820.

Agarwal, Bina. 1994. *A field of one's own: Gender and land rights in South Asia.* Vol. 58 Cambridge University Press.

Agarwal, Bina, Pervesh Anthwal and Malvika Mahesh. 2021. "How many and which women own land in India? Inter-gender and Intra-gender gaps." *The Journal of Development Studies* 57(11):1807–1829.

Agarwal, Bina and Pradeep Panda. 2007. "Toward Freedom from Domestic Violence: The Neglected Obvious." *Journal of Human Development* 8(3):359–388.
URL: *https://doi.org/10.1080/14649880701462171*

Agarwal, Monica. 2024. "The Role of Affirmative Action in Enrollment, Test Scores, and School Quality: Evidence from India.".

Aguirre, Josefa and Juan Matta. 2021. "Walking in your footsteps: Sibling spillovers in higher education choices." *Economics of Education Review* 80:102062.

Aid Water, Unilever Domestos, Water Supply & Sanitation Collaborative Council. 2013. We can't wait: A report on sanitation and hygiene for women and girls. Technical report.
URL: *https://washmatters.wateraid.org/publications/we-cant-wait-a-report-on-sanitation-and-hygiene-for-women-and-girls*

Ajefu, Joseph B, Nadia Singh, Shayequazeenat Ali and Uchenna Efobi. 2022. "Women's inheritance rights and child health outcomes in India." *The Journal of Development Studies* 58(4):752–767.

Alasino, Enrique, Marıa José Ramırez, Mauricio Romero, Norbert Schady and David Uribe. 2023. "Learning Losses during the COVID-19 pandemic: Evidence from Mexico.".

Alasino, Enrique, María José Ramírez, Mauricio Romero, Norbert Schady and David Uribe. 2024. "Learning losses during the COVID-19 pandemic: Evidence from Mexico." *Economics of Education Review* 98:102492.

Ali, Mohammad, Michael Emch, Jean-Paul Donnay, Mohammad Yunus and RB Sack. 2002. "The spatial epidemiology of cholera in an endemic area of Bangladesh." *Social science & medicine* 55(6):1015–1024.

Almond, Douglas and Bhashkar Mazumder. 2013. "Fetal origins and parental responses." *Annu. Rev. Econ.* 5(1):37–56.

Altmejd, Adam, Andrés Barrios-Fernández, Marin Drlje, Joshua Goodman, Michael Hurwitz, Dejan Kovac, Christine Mulhern, Christopher Neilson and Jonathan Smith. 2021. "O brother, where start thou? Sibling spillovers on college and major choice in four countries." *The Quarterly Journal of Economics* 136(3):1831–1886.

Anand, Gautam, Aishwarya Atluri, Lee Crawfurd, Todd Pugatch and Ketki Sheth. 2023. "Improving school management in low and middle income countries: a systematic review.".

Anderson, Siwan and Garance Genicot. 2015. "Suicide and property rights in India." *Journal of Development Economics* 114:64–78.

Andrabi, Tahir, Benjamin Daniels and Jishnu Das. 2021. "Human capital accumulation and disasters: Evidence from the Pakistan earthquake of 2005." *Journal of Human Resources* pp. 0520–10887R1.

Andrabi, Tahir, Natalie Bau, Jishnu Das and Asim Ijaz Khwaja. 2022. Heterogeneity in School Value-Added and the Private Premium. Technical report National Bureau of Economic Research.

Andrabi, Tahir, Natalie Bau, Jishnu Das, Naureen Karachiwalla and Asim Ijaz Khwaja. 2024. "Crowding in private quality: The equilibrium effects of public spending in education." *The Quarterly Journal of Economics* p. qjae014.

Andrew, Alison and Adam Salisbury. 2023. "The educational experiences of Indian children during COVID-19." *Economics of Education Review* 97:102478.

Angrist, Joshua D and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Angrist, Joshua D, Parag A Pathak and Christopher R Walters. 2013. "Explaining charter school effectiveness." *American Economic Journal: Applied Economics* 5(4):1–27.

Angrist, Joshua D, Sarah R Cohodes, Susan M Dynarski, Parag A Pathak and Christopher R Walters. 2016. "Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice." *Journal of Labor Economics* 34(2):275–318.

Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King and Michael Kremer. 2002. "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment." *American Economic Review* 92(5):1535–1558.

Angrist, Joshua, Eric Bettinger and Michael Kremer. 2006. "Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia." *American Economic Review* 96(3):847–862.

Angrist, Joshua and Guido Imbens. 1995. "Identification and estimation of local average treatment effects.".

Angrist, Joshua, Peter Hull and Christopher R Walters. 2023. "Methods for Measuring School Effectiveness." *Handbook of the Economics of Education* .

Angrist, Noam, Peter Bergman and Moitshepi Matsheng. 2020. School's Out: Experimental Evidence on Limiting Learning Loss Using "Low-Tech" in a Pandemic. Technical report National Bureau of Economic Research.

Anukriti, S, Catalina Herrera-Almanza, Mahesh Karra and Rocio Valdebenito. 2022. "Convincing the Mummy-ji: Improving Mother-in-Law Approval of Family Planning in India." *AEA Papers and Proceedings* 112:568–72.

Anukriti, S, Catalina Herrera-Almanza, Praveen K Pathak and Mahesh Karra. 2020. "Curse of the Mummy-ji: The influence of mothers-in-law on women in India." *American Journal of Agricultural Economics* 102(5):1328–1351.

Arcidiacono, Peter and Michael Lovenheim. 2016. "Affirmative action and the quality–fit trade-off." *Journal of Economic Literature* 54(1):3–51.

Ardington, Cally, Gabrielle Wills and Janeli Kotze. 2021. "COVID-19 learning losses: Early grade reading in South Africa." *International Journal of Educational Development* 86:102480.

Arnold, Fred, Minja Kim Choe and Tarun K Roy. 1998. "Son preference, the family-building process and child mortality in India." *Population studies* 52(3):301–315.

Athey, Susan and Guido W Imbens. 2006. "Identification and inference in nonlinear difference-in-differences models." *Econometrica* 74(2):431–497.

Augsburg, Britta, Bansi Malde, Harriet Olorenshaw and Zaki Wahhaj. 2023. "To invest or not to invest in sanitation: The role of intra-household gender differences in perceptions and bargaining power." *Journal of Development Economics* 162:103074.

Augsburg, Britta, Juan P Baquero, Sanghmitra Gautam and Paul Rodriguez-Lesmes. 2023. "Sanitation and marriage markets in India: Evidence from the Total Sanitation Campaign." *Journal of Development Economics* 163:103092.

Azam, Mehtabul, Amiee Chin and Nishith Prakash. 2013. The returns to English-language skills in India. Technical Report 2.

Azevedo, João Pedro, Amer Hasan, Diana Goldemberg, Koen Geven and Syedah Aroob Iqbal. 2021. "Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates." *The World Bank Research Observer* 36(1):1–40.

Bacher-Hicks, Andrew, Joshua Goodman and Christine Mulhern. 2021. "Inequality in household adaptation to schooling shocks: COVID-induced online learning engagement in real time." *Journal of Public Economics* 193:104345.

Badaracco, Nicolás. 2020. "Time Investment Responses of Parents and Students to School Inputs.".

Bagde, Surendrakumar, Dennis Epple and Lowell Taylor. 2016. "Does affirmative action work? Caste, gender, college quality, and academic success in India." *American Economic Review* 106(6):1495–1521.

Bagde, Surendrakumar, Dennis Epple and Lowell Taylor. 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.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer and Edward Miguel. 2016. "Worms at work: Long-run impacts of a child health investment." *The Quarterly Journal of Economics* 131(4):1637–1680.

Bandiera, Oriana, Niklas Buehren, Markus Goldstein and Imran Rasul. 2020. "Do school closures during an epidemic have persistent effects? Evidence from Sierra Leone in the time of Ebola.".

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak and Jeanne Lafortune. 2013. "Marry for what? Caste and mate selection in modern India." *American Economic Journal: Microeconomics* 5(2):33–72.

Bank, World. 2014. Women, Business and the Law 2014. Removing Restrictions to Enhance Gender Equality. Technical report World Bank Report.

Barcellos, Silvia Helena, Leandro S Carvalho and Adriana Lleras-Muney. 2014. "Child gender and parental investments in India: Are boys and girls treated differently?" *American Economic Journal: Applied Economics* 6(1):157–89.

Bates, Karine. 2004. "The Hindu succession act: One law, plural identities." *The Journal of Legal Pluralism and Unofficial Law* 36(50):119–144.

Beam, Emily A, Priya Mukherjee, Laia Navarro-Sola, Junnatul Ferdosh and Md Afzal Hossain Sarwar. 2021. "Take-up, use, and effectiveness of remote technologies." *Endline Report. Innovations for Poverty Action project "Bangladesh COVID-19 Remote Learning Technologies," protocol* 15594.

Becker, Gary S and Nigel Tomes. 1976. "Child endowments and the quantity and quality of children." *Journal of Political Economy* 84(4, Part 2):S143–S162.

Berry, James and Priya Mukherjee. 2016. "Pricing of private education in urban India: Demand, use and impact." *Unpublished manuscript. Ithaca, NY: Cornell University* .

Berry, James, Rebecca Dizon-Ross and Maulik Jagnani. 2020. Not playing favorites: An experiment on parental fairness preferences. Technical report National Bureau of Economic Research.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics* 119(1):249–275.

Bertrand, Marianne, Rema Hanna and Sendhil Mullainathan. 2010. "Affirmative action in education: Evidence from engineering college admissions in India." *Journal of Public Economics* 94(1-2):16–29.

Bhalotra, Sonia, Rachel Brulé and Sanchari Roy. 2020. "Women's inheritance rights reform and the preference for sons in India." *Journal of Development Economics* 146:102275.

Bharadwaj, Prashant, Juan Pedro Eberhard and Christopher A Neilson. 2018. "Health at birth, parental investments, and academic outcomes." *Journal of labor Economics* 36(2):349–394.

Biswas, Shreya, Upasak Das and Prasenjit Sarkhel. 2024. "Duration of exposure to inheritance law in India: Examining the heterogeneous effects on empowerment." *Review of Development Economics* 28(2):777–799.

Black, Dan A and Jeffrey A Smith. 2006. "Estimating the returns to college quality with multiple proxies for quality." *Journal of labor Economics* 24(3):701–728.

Black, Dan A, Joonhwi Joo, Robert LaLonde, Jeffrey A Smith and Evan J Taylor. 2022. "Simple tests for selection: Learning more from instrumental variables." *Labour Economics* 79:102237.

Black, Dan, Jeffrey Smith and Kermit Daniel. 2005. "College quality and wages in the United States." *German Economic Review* 6(3):415–443.

Black, Sandra E, Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth and Marianne Simonsen. 2021. "Sibling spillovers." *The Economic Journal* 131(633):101–128.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South*." *The Quarterly Journal of Economics* 122(1):73–117.
URL: *+ http://dx.doi.org/10.1162/qjec.121.1.73*

Bleemer, Zachary. 2022. "Affirmative action, mismatch, and economic mobility after California's Proposition 209." *The Quarterly Journal of Economics* 137(1):115–160.

Bol, Thijs. 2020. "Inequality in homeschooling during the Corona crisis in the Netherlands. First results from the LISS Panel.".

Bonesrønning, Hans. 2004. "The determinants of parental effort in education production: do parents respond to changes in class size?" *Economics of Education Review* 23(1):1–9.

Borooah, Vani. 2010. "Inequality in health outcomes in India: The role of caste and religion.".

Borusyak, Kirill and Xavier Jaravel. 2018. *Revisiting event study designs.* SSRN Scholarly Paper ID 2826228, Social Science Research Network, Rochester, NY 2018.

Borusyak, Kirill, Xavier Jaravel and Jann Spiess. 2021. "Revisiting event study designs: Robust and efficient estimation." *arXiv preprint arXiv:2108.12419* .

Bose, Nayana and Shreyasee Das. 2021. "Intergenerational effects of improving women's property rights: Evidence from India." *Oxford Development Studies* 49(3):277–290.

Botticini, Maristella and Aloysius Siow. 2003. "Why Dowries?" *The American Economic Review* 93(4):1385–1398.
URL: *http://www.jstor.org/stable/3132295*

Brown, Jennifer, Kripa Ananthpur and Renee Giovarelli. 2002. *Women's access and rights to land in Karnataka.* Rural Development Institute Seattle.

Browning, Martin and Pierre-Andre Chiappori. 1998. "Efficient intra-household allocations: A general characterization and empirical tests." *Econometrica* pp. 1241–1278.

Bruhn, Jesse. 2019. "The consequences of sorting for understanding school quality." *Unpublished working paper). Retrieved from https://1b50402b-a-62cb3a1a-s-sites. googlegroups. com/site/jessebruhn3/jesse_bruhn_jmp. pdf* .

Brule, RE. 2012. "Gender equity and inheritance reform: evidence from rural India." *Unpublished manuscript. See http://rachelbrule. files. wordpress. com/2012/09/brule_paper1_final. pdf* .

Buhl-Wiggers, Julie, Jason T Kerwin, Ricardo Montero de la Piedra, Jeffrey Smith and Rebecca Thornton. 2023. "Reading for Life: Lasting Impacts of a Literacy Intervention in Uganda.".

Callaway, Brantly and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of econometrics* 225(2):200–230.

Calvi, Rossella. 2016. “Why are older women missing in India? the age profile of bargaining power and poverty.” *Journal of Political Economy* 45:58–67.

Calvi, Rossella. 2020. “Why are older women missing in India? The age profile of bargaining power and poverty.” *Journal of Political Economy* 128(7):2453–2501.

Card, David and Alan B Krueger. 2005. “Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas.” *ILR Review* 58(3):416–434.

Carlana, Michela and Eliana La Ferrara. 2021. “Apart but connected: Online tutoring and student outcomes during the COVID-19 pandemic.”.

Cattan, Sarah, Christine Farquharson, Sonya Krutikova, Angus Phimister, Adam Salisbury and Almudena Sevilla. 2021. *Home learning experiences through the COVID-19 pandemic*. Number R195 IFS Report.

Chabrier, Julia, Sarah Cohodes and Philip Oreopoulos. 2016. “What can we learn from charter school lotteries?” *Journal of Economic Perspectives* 30(3):57–84.

Chatterji, Pinka and Yue Li. 2021. “Effects of COVID-19 on school enrollment.” *Economics of Education Review* 83:102128.

Chiappori, Pierre-André. 1988. “Rational household labor supply.” *Econometrica* 56(1):63–90.

Chiappori, Pierre-Andre. 1992. “Collective labor supply and welfare.” *Journal of Political Economy* 100(3):437–467.

Coffey, Diane, Aashish Gupta, Payal Hathi, Nidhi Khurana, Dean Spears, Nikhil Srivastav and Sangita Vyas. 2014. “Revealed preference for open defecation.” *Economic & Political Weekly* 49(38):43.

Coffey, Diane, Dean E Spears and Angus Deaton. 2017. *Where India goes: abandoned toilets, stunted development and the costs of caste.*

Coffey, Diane and Dean Spears. 2017. *Where India Goes: Abandoned Toilets, Stunted Development and the Costs of Caste.* Harper Collins Publishers.

Cohodes, Sarah R, Elizabeth M Setren and Christopher R Walters. 2021. "Can successful schools replicate? Scaling up Boston's charter school sector." *American Economic Journal: Economic Policy* 13(1):138–67.

Crawfurd, Lee, David K Evans, Susannah Hares and Justin Sandefur. 2023. "Live tutoring calls did not improve learning during the COVID-19 pandemic in Sierra Leone." *Journal of Development Economics* 164:103114.

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

Cunha, Flavio, James J Heckman and Susanne M Schennach. 2010. "Estimating the technology of cognitive and noncognitive skill formation." *Econometrica* 78(3):883–931.

Damera, Vijay Kumar. 2018. Essays on school choice PhD thesis University of Oxford.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan and Venkatesh Sundararaman. 2013. "School inputs, household substitution, and test scores." *American Economic Journal: Applied Economics* 5(2):29–57.

Das, Kumudin, KC Das, TK Roy and PK Tripathy. 2011. "Dynamics of inter-religious and inter-caste marriages in India." *Population Association of America, Washington DC, USA* .

Datar, Ashlesha and Bryce Mason. 2008. "Do reductions in class size "crowd out" parental investment in education?" *Economics of Education Review* 27(6):712–723.

De Chaisemartin, Clément and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review* 110(9):2964–2996.

Deininger, Klaus, Aparajita Goyal and Hari Nagarajan. 2013. "Women's inheritance rights and intergenerational transmission of resources in India." *Journal of Human Resources* 48(1):114–141.

Deininger, Klaus, Songqing Jin, Hari K Nagarajan and Fang Xia. 2019. "Inheritance law reform, empowerment, and human capital accumulation: Second-generation effects from India." *The Journal of Development Studies* 55(12):2549–2571.

Del Boca, Daniela, Chiara Monfardini and Cheti Nicoletti. 2017. "Parental and child time investments and the cognitive development of adolescents." *Journal of Labor Economics* 35(2):565–608.

Deming, David J, Justine S Hastings, Thomas J Kane and Douglas O Staiger. 2014. "School choice, school quality, and postsecondary attainment." *American Economic Review* 104(3):991–1013.

Desai, Sonalde B, Amaresh Dubey, Brij Lal Joshi, Mitali Sen, Abusaleh Shariff and Reeve Vanneman. 2010. "Human development in India." *New York: Oxford University* .

Dessy, Sylvain, Horace Gninafon, Luca Tiberti and Marco Tiberti. 2023. "Free compulsory education can mitigate COVID-19 disruptions' adverse effects on child schooling." *Economics of Education Review* 97:102480.

Dillon, Eleanor Wiske and Jeffrey Andrew Smith. 2020. "The consequences of academic match between students and colleges." *Journal of Human Resources* 55(3):767–808.

Dobbie, Will and Roland G Fryer Jr. 2011. "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3(3):158–187.

Dobbie, Will and Roland G Fryer Jr. 2013. "Getting beneath the veil of effective schools: Evidence from New York City." *American Economic Journal: Applied Economics* 5(4):28–60.

Dobbie, Will, Roland G Fryer et al. 2011. Exam high schools and academic achievement: Evidence from New York City. Technical report National Bureau of Economic Research.

Duflo, Esther. 2003. "Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa." *The World Bank Economic Review* 17(1):1–25.

Engzell, Per, Arun Frey and Mark D Verhagen. 2021. "Learning loss due to school closures during the COVID-19 pandemic." *Proceedings of the National Academy of Sciences* 118(17):e2022376118.

Fredriksson, Peter, Björn Öckert and Hessel Oosterbeek. 2016. "Parental responses to public investments in children: Evidence from a maximum class size rule." *Journal of Human Resources* 51(4):832–868.

Gelber, Alexander and Adam Isen. 2013. "Children's schooling and parents' behavior: Evidence from the Head Start Impact Study." *Journal of Public Economics* 101:25–38.

Geruso, Michael and Dean Spears. 2018. "Neighborhood sanitation and infant mortality." *American Economic Journal: Applied Economics* 10(2):125–62.

Glewwe, Paul and Karthik Muralidharan. 2016. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*. Vol. 5 Elsevier pp. 653–743.

Glewwe, Paul and Michael Kremer. 2006. "Schools, teachers, and education outcomes in developing countries." *Handbook of the Economics of Education* 2:945–1017.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225(2):254–277.

Goolsbee, Austan and Amil Petrin. 2004. "The consumer gains from direct broadcast satellites and the competition with cable TV." *Econometrica* 72(2):351–381.

Greaves, Ellen, Iftikhar Hussain, Birgitta Rabe and Imran Rasul. 2019. Parental responses to information about school quality: Evidence from linked survey and administrative data. Technical report ISER Working Paper Series.

Greaves, Ellen, Iftikhar Hussain, Birgitta Rabe and Imran Rasul. 2023. "Parental responses to information about school quality: Evidence from linked survey and administrative data." *The Economic Journal* 133(654):2334–2402.

Guariso, Andrea and Martina Björkman Nyqvist. 2023. The Impact of the COVID-19 pandemic on children's learning and wellbeing: Evidence from India. Technical report Stockholm School of Economics, Mistra Center for Sustainable Markets (Misum).

Guo, Naijia, Shuangxin Wang and Junsen Zhang. 2024. "The Short-and Long-Run Impacts of Free Education on Schooling: Direct Effects and Intra-Household Spillovers." *The Economic Journal* p. ueae049.

Gupta, Sweta, Christopher Ksoll and Annemie Maertens. 2021. "Intra-household efficiency in extended family households: Evidence from rural India." *The Journal of Development Studies* 57(7):1172–1197.

Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *The Economic Journal* 113(485):F64–F98.

Harbatkin, Erica, Katharine O Strunk and Aliyah McIlwain. 2023. "School turnaround in a pandemic: An examination of the outsized implications of COVID-19 on low-performing turnaround schools, districts, and their communities." *Economics of Education Review* 97:102484.

Hassan, Hashibul, Asad Islam, Abu Siddique, Liang Choon Wang et al. 2021. Telementoring and homeschooling during school closures: A randomized experiment in rural Bangladesh. Technical report TUM School of Governance at the Technical University of Munich.

Hathi, Payal, Sabrina Haque, Lovey Pant, Diane Coffey and Dean Spears. 2014. "Place and child health: the interaction of population density and sanitation in developing countries.".

Heath, Rachel and Xu Tan. 2014. "Intrahousehold bargaining, female autonomy, and labor supply: Theory and evidence from India." *University of Washington* .

Heckman, James, Jeffrey Smith and Christopher Taber. 1998. "Accounting for dropouts in evaluations of social programs." *Review of Economics and Statistics* 80(1):1–14.

Hidrobo, Melissa, Jessica B Hoel and Katie Wilson. 2021. "Efficiency and status in polygynous pastoralist households." *The Journal of Development Studies* 57(2):326–342.

Hossain, Md Amzad, Kanika Mahajan and Sheetal Sekhri. 2022. "Access to toilets and violence against women." *Journal of Environmental Economics and Management* 114:102695.

Houtenville, Andrew J and Karen Smith Conway. 2008. "Parental effort, school resources, and student achievement." *Journal of Human resources* 43(2):437–453.

Hsieh, Chang-Tai and Miguel Urquiola. 2006. "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program." *Journal of public Economics* 90(8-9):1477–1503.

Imbens, Guido W and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467–475.

Indus Action. 2019. The Bright Spots Report: Status of social inclusion through RTE Section 12(1)(c). Technical report Indus Action.

Jack, Rebecca, Clare Halloran, James Okun and Emily Oster. 2023. "Pandemic schooling mode and student test scores: Evidence from US school districts." *American Economic Review: Insights* 5(2):173–190.

Jackson, Kirabo. 2010. "Do students benefit from attending better schools? Evidence from rule-based student assignments in Trinidad and Tobago." *The Economic Journal* 120(549):1399–1429.

Jackson, Kirabo, Rucker C Johnson and Claudia Persico. 2016. "The effects of school spending on educational and economic outcomes: Evidence from school finance reforms." *The Quarterly Journal of Economics* 131(1):157–218.

Jadhav, Apoorva, Abigail Weitzman and Emily Smith-Greenaway. 2016. "Household sanitation facilities and women's risk of non-partner sexual violence in India." *BMC public health* 16(1):1139.

Jayachandran, Seema and Ilyana Kuziemko. 2011. "Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India." *The Quarterly Journal of Economics* 126(3):1485–1538.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam and Rahim Moineddin. 2006. "Low male-to-female sex ratio of children born in India: national survey of 1 · 1 million households." *The Lancet* 367(9506):211–218.

Joensen, Juanna Schrøter and Helena Skyt Nielsen. 2018. "Spillovers in education choice." *Journal of Public Economics* 157:158–183.

Kelkar, Govind. 2014. "The fog of entitlement: Women's inheritance and land rights." *Economic and Political Weekly* pp. 51–58.

Khanna, Gaurav. 2020. "Does affirmative action incentivize schooling? Evidence from India." *Review of Economics and Statistics* 102(2):219–233.

Khanna, Gaurav. 2023. "Large-scale education reform in general equilibrium: Regression discontinuity evidence from India." *Journal of Political Economy* 131(2):549–591.

Khanna, Tina and Madhumita Das. 2016. "Why gender matters in the solution towards safe sanitation? Reflections from rural India." *Global public health* 11(10):1185–1201.

Kilenthong, Weerachart T, Khanista Boonsanong, Sartja Duangchaiyoosook, Wasinee Jantorn and Varunee Khruapradit. 2023. "Learning losses from school closure due to the COVID-19 pandemic for Thai kindergartners." *Economics of Education Review* 96:102455.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing incomplete political science data: An alternative algorithm for multiple imputation." *American Political Science Review* 95(1):49–69.

Kingdon, Geeta Gandhi. 2007. "The progress of school education in India." *Oxford Review of Economic Policy* 23(2):168–195.

Kingdon, Geeta Gandhi. 2020. "The private schooling phenomenon in India: A review." *The Journal of Development Studies* 56(10):1795–1817.

Kishore, Avinash and Dean Spears. 2012. Clean cooking fuel, women's intrahousehold status, and son preference in rural India. In *Poster presentation, Population Association of America annual meeting, San Francisco.*

Kline, Patrick, Evan Rose and Christopher Walters. 2022. Systemic discrimination among large US employers. Technical Report 4.

Kuhfeld, Megan, James Soland, Beth Tarasawa, Angela Johnson, Erik Ruzek and Jing Liu. 2020. "Projecting the potential impact of COVID-19 school closures on academic achievement." *Educational Researcher* 49(8):549–565.

Lai, Fang, Elisabeth Sadoulet and Alain de Janvry. 2011. "The contributions of school quality and teacher qualifications to student performance evidence from a natural experiment in Beijing middle schools." *Journal of Human Resources* 46(1):123–153.

Landerso, Rasmus, Helena Skyt Nielsen and Marianne Simonsen. 2017. "How going to school affects the family." *Department of Economics Aarhus University* .

Maldonado, Joana Elisa and Kristof De Witte. 2022. "The effect of school closures on standardized student test outcomes." *British Educational Research Journal* 48(1):49–94.

Ministry of Women and Child Development. 2020. "Operations of Anganwadi services during COVID-19.". Retrieved from https://wcd.nic.in/sites/default/files/AWC%20services%20continuation_0.pdf.

Mookerjee, Sulagna. 2019. "Gender-neutral inheritance laws, family structure, and women's status in India." *The World Bank Economic Review* 33(2):498–515.

Moscoviz, Laura, David K Evans et al. 2022. *Learning loss and student dropouts during the COVID-19 pandemic: A review of the evidence two years after schools shut down.* Center for Global Development.

Muralidharan, Karthik and Michael Kremer. 2006. "Public and private schools in rural India." *Harvard University, Department of Economics, Cambridge, MA* 9:10–11.

Muralidharan, Karthik and Venkatesh Sundararaman. 2015. "The aggregate effect of school choice: Evidence from a two-stage experiment in India." *The Quarterly Journal of Economics* 130(3):1011–1066.

Musaddiq, Tareena, Kevin Stange, Andrew Bacher-Hicks and Joshua Goodman. 2022. "The pandemic's effect on demand for public schools, homeschooling, and private schools." *Journal of Public Economics* 212:104710.

Nallari, Anupama. 2015. ““All we want are toilets inside our homes!”: The critical role of sanitation in the lives of urban poor adolescent girls in Bengaluru, India.” *Environment and Urbanization* 27(1):73–88.
URL: *https://doi.org/10.1177/0956247814563514*

Nicoletti, Cheti and Birgitta Rabe. 2019. “Sibling spillover effects in school achievement.” *Journal of Applied Econometrics* 34(4):482–501.

Opper, Isaac M. 2019. “Does helping John help Sue? Evidence of spillovers in education.” *American Economic Review* 109(3):1080–1115.

Patrinos, Harry Anthony, Emiliana Vegas and Rohan Carter-Rau. 2022. “An analysis of COVID-19 student learning loss.”.

Pop-Eleches, Cristian and Miguel Urquiola. 2013. “Going to a better school: Effects and behavioral responses.” *American Economic Review* 103(4):1289–1324.

Pratham. 2019. “Annual Status of Education Report. Technical report Pratham.

Qian, Nancy. 2008. “Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance.” *The Quarterly Journal of Economics* 123(3):1251–1285.

Rabe, Birgitta. 2020. “Schooling inputs and behavioral responses by families.” *Handbook of Education Economics: A Comprehensive Overview* pp. chap. 16, 217—-227, 2nd ed.

Rao, Gautam. 2019. “Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools.” *American Economic Review* 109(3):774–809.

Romero, Mauricio and Abhijeet Singh. 2022. The incidence of affirmative action: Evidence from quotas in private schools in India. Technical report Working paper.

Romero, Mauricio and Abhijeet Singh. 2023. “The incidence and effects of affirmative action: Evidence from quotas in private schools in India.”.

Romero, Mauricio, Justin Sandefur and Wayne Aaron Sandholtz. 2020. "Outsourcing education: Experimental evidence from Liberia." *American Economic Review* 110(2):364–400.

Rosenblum, Daniel. 2015. "Unintended consequences of women's inheritance rights on female mortality in India." *Economic Development and Cultural Change* 63(2):223–248.

Rosenzweig, Mark R and T Paul Schultz. 1982. "Market opportunities, genetic endowments, and intrafamily resource distribution: Child survival in rural India." *The American Economic Review* 72(4):803–815.

Roth, Jonathan. 2013. Interpreting Event-Studies from Recent Difference-in-Differences Methods. Technical report.

Roy, Sanchari. 2008. "Female empowerment through inheritance rights: evidence from India." *Mimeo, London School of Economics, London* .

Roy, Sanchari. 2015. "Empowering women? Inheritance rights, female education and dowry payments in India." *Journal of Development Economics* 114:233–251.

Sant'Anna, Pedro HC and Jun Zhao. 2020. "Doubly robust difference-in-differences estimators." *Journal of Econometrics* 219(1):101–122.

Sen, Amartya. 2003. "Missing women revisited: reduction in female mortality has been counterbalanced by sex selective abortions." *BMJ: British Medical Journal* 327(7427):1297.

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

Singh, Abhijeet, Mauricio Romero and Karthik Muralidharan. 2022. COVID-19 Learning loss and recovery: Panel data evidence from India. Technical report National Bureau of Economic Research.

Singh, Abhijeet, Mauricio Romero and Karthik Muralidharan. 2024. "COVID-19 Learning loss and recovery: Panel data evidence from India." *Journal of Human Resources* .

Stopnitzky, Yaniv. 2017. “No toilet no bride? Intrahousehold bargaining in male-skewed marriage markets in India.” *Journal of Development Economics* 127:269–282.

Strauss, John, Germano Mwabu and Kathleen Beegle. 2000. “Intrahousehold allocations: a review of theories and empirical evidence.” *Journal of African Economies* 9(Supplement_1):83–143.

Thomas, Duncan. 1990. “Intra-household resource allocation: An inferential approach.” *Journal of Human Resources* pp. 635–664.

Tisdell, C.A and K.C Roy. 2002. “Property rights in women’s empowerment in Rural India: A review.” *International Journal of Social Economics* 29(4):315–334.

Todd, Petra E and Kenneth I Wolpin. 2003. “On the specification and estimation of the production function for cognitive achievement.” *The Economic Journal* 113(485):F3–F33.

Tooley, James. 2013. *The beautiful tree: A personal journey into how the world’s poorest people are educating themselves.* Cato Institute.

Tooley, James and Pauline Dixon. 2007. “Private education for low-income families: Results from a global research project.” *Private schooling in less economically developed countries: Asian and African perspectives* pp. 15–39.

Weiland, Christina, Rebecca Unterman, Susan Dynarski, Rachel Abenavoli, Howard Bloom, Breno Braga, Anne-Marie Faria, Erica Greenberg, Brian Jacob, Jane Arnold Lincove et al. 2023. “Lottery-Based Evaluations of Early Education Programs: Opportunities and Challenges for Building the Next Generation of Evidence.”.

Wolf, Sharon, Elisabetta Aurino, Noelle M Suntheimer, Esinam A Avornyo, Edward Tsinigo, Jasmine Jordan, Solomon Samanhiya, J Lawrence Aber and Jere R Behrman. 2022. “Remote learning engagement and learning outcomes during school closures in Ghana.” *International Journal of Educational Research* 115:102055.

World Bank. 2021. “Remote learning during the global school lockdown: Multi-country lessons.”.

WSP. 2011. A Decade of the Total Sanitation Campaign: Rapid Assessment of Processes and Outcomes. Technical report.