Essays on the Economics of Education in Developing Countries

Lee Crawfurd

Submitted for the degree of Doctor of Philosophy

Department of Economics

University of Sussex

October 2019

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree. An abridged version of Chapter One has been published in the *Journal of African Economies*. Chapter Two is co-authored with Dev Patel and Justin Sandefur.

Signature: Lee Crawfurd

UNIVERSITY OF SUSSEX

Lee Crawfurd, Doctor of Philosophy

Essays on the Economics of Education in Developing Countries

SUMMARY

This thesis is comprised of three essays on the economics of education in developing countries, focused on the analysis of public financing of schools operated by non-state actors.

Chapter One focuses on publicly subsidised private secondary schools in Uganda. Student value-added is higher in these schools than in government schools. The chapter explores the role of the quality of school management as a mediating factor in the performance differential, finding that private schools are on average no better managed than government schools, with the exception of those managed by an international charity.

Chapter Two evaluates a five-year private school voucher lottery programme in Delhi. This lottery was designed as a test of India's national Right to Education Act Section 12(1)(c), which reserves 25 percent of places at private schools nationwide for students from "economically weaker sections", with funding coming from government. Students who won the lottery and attended low-cost private schools performed slightly worse in Hindi and no different in Maths and English, or on various non-cognitive skills.

Chapter Three evaluates a large-scale contracting out of public schools to private management in Punjab, Pakistan. Using a difference-in-difference framework, I estimate that failing government schools that are contracted out to private operators dramatically increase their enrolment, but that the effect on student learning is ambiguous.

Overall these three studies highlight the variability in forms of public financing for independently operated schools, and the variability in quality. In all three policies, financing for non-state schools costs significantly less than equivalent spending in government schools.

Acknowledgements

I have a lot of people to thank for their advice and support in the preparation of this thesis. First of all Abhijeet Singh for persuading me to actually start this thing in the first place. I think it was the right decision, but I can't observe the counterfactual. Anthony Harris, Gregory Smith, and Gunilla Pettersson all helped to convince me that it was possible to finish and live to tell the tale.

I am grateful to my advisors Andy McKay and Iftikhar Hussain, and to others at Sussex including Peter Dolton, Shqiponja Telhaj, Giordano Mion, and my PhD cohort for helpful comments, including from Rashaad Shabab, Panka Bencsik, Gustavo Iriarte Rivas, Antonia Shwarz, and Elsa Valli.

I benefited enormously from time spent at the Center for Global Development and Ark's Education Partnerships Group. Matthew Collin and Theodore Talbot provided excellent advice and distraction. Leaders at the Research on Improving Systems of Education (RISE) programme provided substantial inspiration, namely Lant Pritchett, Justin Sandefur, Clare Leaver, Amanda Beatty, and Caine Rolleston. At Ark, Susannah Hares backed surveys in Uganda and India, and Phil Elks and Sabina Morley pre-tested and helped to adapt the World Management Survey (WMS) schools instrument for Uganda. Phil Elks also recruited a fantastic team of interviewers.

I'm grateful to my co-authors on the India chapter Dev Patel and Justin Sandefur. I'm indebted to Venkitesan and Kruti Bharucha at Ark India, and Arshi Naaz and her team from Srajan Shiksha Foundation for superb data collection. I am grateful to Abdullah Alam for his advice on Pakistan, and Laura Moscoviz for research assistance.

I have received helpful comments on different chapters from Jacobus Cilliers, Jishnu Das, Edward Davis, Ahmad Jawad Asghar, Nadia Naviwala, Rahila Fatima, Imtiaz Dossa, Maryam Akmal, and conference participants at the Centre for the Study of African Economies in Oxford, Royal Economic Society, RISE, North-Eastern Universities Development Conference, Goettingen International Labour and Development, and from two anonymous referees at the Journal of African Economies.

I am grateful for financial support from the Department for International Development - funded RISE programme and the Economic and Social Research Council Doctoral Training Grant.

The Uganda National Examinations Board, the Punjab Examinations Commission, Punjab Education Foundation, and Ark Education Partnerships Group generously provided data used in this thesis. I'm grateful to all of the headteachers, parents, and children who gave their time to be interviewed.

Finally the biggest thanks of all to my family – to Jenny for priceless investments in my own human capital, to Tina Turner for countless hours of childcare, and Georgina, Rowan, and Joshua for all of your love and incredible patience.

Contents

List of Tables

List of Figures

Abbreviations

ASER	Annual Status of Education Report
CEC	Colegios en Concesión
DISE	District Information System for Education
EDO	Executive District Officer
EMIS	Education Management Information System
ENABLE	Ensuring Access to Better Learning Experiences
EVS	Education Voucher Scheme
FAS	Foundation Assisted Schools
GDP	Gross Domestic Product
GPCM	Generalized Partial Credit Model
IRT	Item Response Theory
LND	Literacy and Numeracy Drive
LPM	Linear Probability Model
MEA	Monitoring and Evaluation Assistant
NGO	Non Governmental Organisation
NSP	New Schools Programme
NTC	National Teacher College
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
ORF	Oral Reading Fluency
PCA	Principal Component Analysis
PEC	Punjab Examinations Commission
PEF	Punjab Education Foundation
PEIMA	Punjab Education Initiative Management Authority

PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PLE	Primary Leaving Exam
PMIU	Programme Monitoring and Implementation Unit
PPP	Public-Private Partnerships
PSL	Partnership Schools for Liberia
PSM	Propensity Score Matching
PSSP	Public School Support Programme
QAT	Quality Assurance Test
RCT	Randomised Control Trial
RD	Regression Discontinuity
RTE	Right To Education
SD	Standard Deviation
SED	School Education Department
UCE	Uganda Certificate of Education
UK	United Kingdom
UNEB	Uganda National Examinations Board
UNESCO	United Nations Educational, Scientific and Cultural Organisation
USD	United States Dollars
USE	Universal Secondary Education programme
WDI	World Development Indicators
WMS	World Management Survey

Introduction

There have been rapid increases in the number of children attending school around the world, but there remains a global crisis of learning in schools. Of around half a billion children attending school worldwide, an estimated 130 million haven't learnt basic skills even after attending school for four or more years (UNESCO, 2014). Fewer than one in five upper primary school students in Sub-Saharan Africa can meet a global minimum standard of proficiency in mathematics or reading. The average grade 6 student in New Delhi, India, performed at a grade 3 level in Mathematics in 2015 (World Bank, 2018).

A key cause of such weak results has been failures in governance. A basic condition for learning to take place is for teachers to be present. A survey of seven African countries (Kenya, Mozambique, Nigeria, Senegal, Tanzania, Togo, and Uganda) found that in unannounced visits 44 percent of teachers were absent from class (Bold et al., 2017). In one third of classrooms there were students present with no teacher. Combining teacher presence estimates with data on how lesson time was used when there was a teacher, (Bold et al., 2017) estimate that students are actually taught for an average of 2 hours and 46 minutes per day, roughly half of the scheduled time.

Alongside such weak results and governance failures in the public school sector, there has been an explosion in the low-cost private school sector, particularly in middle-income countries. The share of children attending private schools in middle-income countries almost doubled from around 10 percent in 2000 to 19 percent in 2015.[1] In response to this growth in demand for private schools and failures in governance in the public sector, many governments have explored ways to involve the private sector in the delivery of public schooling. The theory is that by increasing accountability for outcomes in exchange for more operational autonomy, schools will be able to deliver better results. The nature of teaching means local knowledge is necessary, and by providing autonomy on implementation (but not on standards or evaluation) to local actors, better performance may be able to be achieved (Pritchett, 2013).

In this thesis I evaluate three different policies involving public financing for non-government operated schools in three different developing countries. Aslam et al. (2017) offer a typology of three different types of public-private partnership in education; subsidies for private schools, vouchers for children to attend private schools, and government schools that are contracted to private operators (dubbed contract schools). Subsidies and voucher schemes both involve government financing for students to attend existing private schools. The main conceptual distinction between these two types is whether the money is targeted at the school or the individual student, though in practice in many schemes labelled as "voucher" programmes there is no actual physical voucher and financing is provided directly to the school. Whilst critics of the privatisation of public schools abound, some of these schemes are perhaps actually closer to "public-ising private schools"[2]. The category of "contract schools" includes well-known charter school schemes in the United States and academy schools in the United Kingdom.

[1] UNESCO Institute for Statistics

[2] https://www.huffingtonpost.com/entry/punjab-and-sindh-provinces-in-pakistan-are-public-ising_us_59e474eee4b02e99c5835804

[2] https://www.huffingtonpost.com/entry/punjab-and-sindh-provinces-in-pakistan-are-public-ising_us_59e474eee4b02e99c5835804

In Chapter One I focus on government subsidised private schools in Uganda. A common perception is that private schools achieve better results than government schools through better management practices. I test this theory directly, presenting the first internationally benchmarked estimates of school management quality from Africa (based on the "World Management Survey"), and how much this mediates the effect of school type on student outcomes. I thus contribute to literatures on the role of school governance, organization, and management in determining student outcomes. I find that student value-added is higher in non-state schools than in government schools, after controlling for household income. Despite this better performance, I find that private schools are on average no better managed than government schools (with one exception being those managed by an international charity). The overall level and distribution of management quality is similar to that found in other non high-income countries (India and Brazil). Despite the result that management quality does not explain the better performance of non-state schools, I show that overall differences in school management quality do matter for student value-added - a standard deviation change in management is associated with a 0.06 standard deviation change in test scores.

Chapter Two evaluates a five-year private school voucher lottery programme in Delhi. This lottery was designed as a test of India's national Right to Education Act Section 12(1)(c), which reserves 25 percent of places at private schools nationwide for students from "economically weaker sections", with funding coming from government. Students who won the lottery were provided with vouchers to cover the cost of attending a low-cost private school in East Delhi. Lottery winners (and students induced to attend a private school) performed slightly *worse* in Hindi and no different in Maths and English, or on various non-cognitive skills. By focusing on urban schools, this chapter complements existing research on private schools in India that primarily looks at rural areas. It also adds to the global evidence base on the quality of low-cost private schools in developing countries. We explore three potential explanations for the puzzle of why so many parents pay for schools that are in fact no better than government schools, finding some support for heterogeneous effects and for asymmetric information.

Chapter Three evaluates a large-scale contracting out of public schools to private management in Punjab, Pakistan. 4,276 weak government primary schools (around 12 percent of the total) were contracted out to private operators in a single year. These schools remain free to students and the private operator receives a per-student subsidy equivalent to less than half of per-pupil spending in government schools. Using a difference-in-difference framework, I estimate that enrolment in converted schools increased by over 60 percent. Converted schools see a slight decline in overall average test scores, but we are unable to distinguish if this is a compositional effect or a treatment effect. Schools with no increase in the number of students sitting the test saw no change in average test scores. This is the first study to estimate the effect of a large scale charter-style contracting of public schools to non-state providers in a developing country (just four prior studies of such programmes have considered either well-identified effects of relatively small scale programmes or descriptive analysis of the precursor to this large-scale programme in Pakistan).

These three Chapters show some learning gains from some kinds of public—private partnerships in some contexts, but not all. Together they highlight the variability in forms of public financing for independently operated schools, and the variability in quality of those schools. These chapters add to the limited global evidence-base on such school reforms, which remains too thin and too mixed to be able to draw strong policy conclusions. One important commonality is that in all three policies, financing for non-state schools costs significantly less than equivalent spending in government schools. This cost differential is driven by lower market salaries than civil service salaries for teachers in each of the three contexts. In two of the cases the policy was at least successful in cost-effectively increasing school enrolment, but in none was it an effective solution to the learning crisis. Providing schools with autonomy in exchange for accountability may be part of the solution, but it probably won't be sufficient by itself, at least on any reasonable time scale.

Chapter 1: School Management and Public–Private Partnerships in Uganda

1.1 Introduction

Can the quality of school management explain differences in student test scores? School productivity varies substantially both within and between countries, and this matters. Theory and evidence suggest that it is the skills and knowledge acquired that lead to higher earnings, not just the amount of time spent in school (Hanushek 2013; Hanushek et al. 2015). In this paper I first ask how much the quality of school management matters for student outcomes. I then consider whether differences in management quality explain differences in the performance of government and private schools, and finally look at what factors explain variation in management quality. I find that school management quality does indeed matter for productivity, as measured by student value-added. However there are no differences on average in management quality between government and private schools, leaving the private school premium unexplained. An important exception is a UK-owned chain of public-private partnership (PPP schools), which are substantially better managed than average, and this difference in management quality explains their performance advantage. Few other factors reliably predict management quality.

This paper is connected to three sets of literature. First, the literature on school effectiveness has sought to identify what makes schools more or less productive in terms of student learning. The landmark Coleman report highlighted how the majority of variation in student achievement is due to family background rather than observable school characteristics (Coleman et al., 1966). Since then hundreds of studies have explored the relationships between educational inputs and school productivity. conduct a comprehensive review of studies from developing countries between 1990 and 2010 that estimate the effect of school or teacher characteristics on student outcomes. They find 79 studies, of which the majority find no statistically significant relationship. One more promising avenue has been estimating the total effect of individual teachers on learning with value-added models. Teachers matter a great deal (Hanushek & Rivkin, 2010). Turning to interventions designed to improve school quality in developing countries, the most successful have focused on pedagogy and governance (Glewwe & Muralidharan, 2015). This includes studies looking at bundled packages of management support (Fryer 2017; Blimpo et al., 2015; Tavares 2015; Lassibille, 2016; (Beasley & Huillery, 2017), as well as studies focused on specific sub-components of school management, such as monitoring (de Hoyos, Garcia-Moreno, & Patrinos, 2015), teacher management (Duflo, Dupas, and Kremer 2015; Muralidharan and Sundararaman 2011; Atherton and Kingdon 2010), and tailoring teaching to the right level of individual students (Pritchett and Beatty 2015; Duflo, Dupas, and Kremer 2011; Banerjee et al., 2016).

A second set of literature looks at the “New Empirical Economics of Management”, demonstrating links between new measures of management practices and productivity in a variety of sectors, including manufacturing, retail, healthcare, and education (Bloom, Lemos, Sadun, Scur, & Reenen, 2014). Better managed manufacturing firms have higher levels of sales, sales growth, profitability, and a lower chance of exit (Bloom, Sadun, & Reenen, 2012). Management quality can explain the productivity gap between US multinationals in Europe and non-US multinationals (Bloom, Sadun, et al., 2012). Better managed hospitals have lower mortality rates, and this measure of management quality responds positively to competition (Bloom, Propper, Seiler, & Van Reenen, 2016). In schools there is a positive correlation between measured management quality and school average test scores in seven different countries (Bloom, Lemos, Sadun, & Van Reenen, 2015).

Third, several papers seek to identify the sources of differences in productivity between regular government schools and schools given increased autonomy under different public-private partnership arrangements, known as Charter schools in the US and Academies in the UK ((Dobbie & Fryer, 2013); Angrist, Pathak, and Walters 2013; Eyles, Hupkau, and Machin 2016). Whilst these studies do suggest that providing operational autonomy to schools can improve performance, this is within a context of a broadly functional education system that provides clear objectives and accountability for schools. Greater autonomy may not produce the same results if schools are not held accountable for their performance. Cross-country studies looking at changes in the level of school autonomy have found that increases in school autonomy lead to better performance in high income countries but worse performance in low-income countries (Hanushek et al. 2013; Contreras 2015). Whilst "autonomous government schools" seem to be better managed than average in OECD countries, there are no comparable estimates from developing countries[3].

In this paper I provide the first estimates of school management quality from sub-Saharan Africa using a version of the "World Management Survey". Schools in

---

[3] Bloom et al (2015) do look at private aided schools in India, finding them no better performing than regular government schools. This should not be surprising however, as these schools have much less autonomy than charter schools or academies. Their teachers are recruited and paid by a central government Education Service Commission rather than by the school, and resemble regular government schools much more than private schools (Kingdon, 2007).

my nationally representative sample of Ugandan secondary schools score on average 2.0 points on a one to five scale, placing them above India and slightly below Brazil. I then demonstrate that management quality matters for results for student performance in high-stakes tests, using a lagged dependent variable dynamic OLS value-added framework, controlling for student prior attainment and school characteristics. This marks a methodological improvement upon previous work that looks at raw correlations between management quality and school average test scores.

I find that despite having more autonomy, private schools and PPP schools are no better managed than government schools. An exception is a chain of international PPP schools (run by the UK education NGO 'PEAS') that have a strong internal performance management framework with high stakes for head teachers. These schools are substantially (more than 1 standard deviation) better managed than average, and perform commensurately better in terms of student value-added.

Finally, conceiving of management as a technology (Bloom, Sadun, & Van Reenen, 2016), I contribute to the literature on technology adoption in developing countries, looking at what factors correlate with better management practices. Schools with greater autonomy and in geographical areas with a greater supply of skilled workers have better management practices, but other headteacher and school factors are not correlated with better practice.

1.2 The School System in Uganda

Uganda introduced free universal primary schooling in 1997, and free secondary schooling in 2007. Enrolment rates have risen accordingly - the net enrolment rate at primary level is now above 90 percent, but the primary completion rate is only around 54 percent, and the secondary rates are lower; around 23 percent net enrolment and 29 percent junior secondary completion (The World Bank, 2013). The official age of school entry is six years old (median age currently in the first grade is seven years old in 2012 survey data[4]), and there are seven grades of primary school (P1 - P7), followed by four years of lower secondary (S1 - S4) and two years of senior secondary (S5 - S6).

The Uganda National Examinations Board (UNEB) administers exams at the end of the last year of primary school (P7) to pupils in both public and private schools (the Primary Leaving Exam or PLE). It is a requirement to pass this exam in order to progress to secondary school. Of 627,000 students enrolled in P7 in 2014, 586,000 (93 percent) registered and sat the PLE, and 517,000 (82 percent) passed. Students take exams in four subjects; English, Maths, Science, and Social Studies. Within each subject a score is given between one and nine, in which a score of 1 - 2 is a Distinction, 3 - 6 is a Credit, 7 - 8 is a Pass, and 9 a Fail. UNEB reported 909 cases of exam malpractice in 2015 (cheating by collusion, external assistance or impersonation), down from 1,344 cases in 2014.

At secondary school, Ordinary level exams – the Uganda Certificate of Education (UCE) - are taken after four years in a minimum of eight subjects, and Advanced

---

[4] 2011-12 Third wave of the Uganda National Panel Survey

level exams taken after two further years in three subjects. The UCE comprises six mandatory subjects administered in English; these are Mathematics, English language, Biology, Chemistry, Physics, and a choice of either Geography, or History, or Religious education. The two final optional subjects can include cultural subjects (such as Music); technical subjects (such as Carpentry); or other subjects such as Accounting, Business and Computer science. As for PLE scores, UCE scores are given for each subject between one and nine, where a 1 - 2 is a Distinction and a 9 is a Fail.

There are 1,007 government secondary schools and 1,785 private secondary schools, of which some from both sector are part of the free Universal Secondary Education programme (USE). Government schools have on average nearly twice as many students as private schools. Table 1.1 shows summary statistics.

Teacher recruitment is managed centrally for public schools. Schools submit vacancies to the Ministry of Education, who then allocate teachers to schools. The Ministry of Public Service pays teachers directly into their bank account, making it difficult for schools to vary pay according to performance. In the private sector teachers are paid substantially lower wages and schools follow their own recruitment procedures (Ugandan Ministry of Education and Sport & UNESCO - IIEP Pôle de Dakar, 2014). Government teacher starting salaries according to the Public Service Pay scale are 511,000 Ugandan Shillings (UGX) ($150 USD) per month. Data is not available for private sector teachers, but across all occupations, median monthly wages were 330,000 UGX ($100 USD) in the public sector and 99,000 UGX ($30 USD) in the private sector for those in paid employment aged 14 - 64 in the 2012/13 National Household Survey (Uganda Bureau of Statistics, 2014).

Free Universal Secondary Education (USE) Programme

The Universal Secondary Education (USE) programme offers free places at registered schools for eligible pupils. Most government schools are registered for USE, with the exception of a minority of elite schools that opt out. Due to the limited number of government secondary schools when the programme was introduced, private schools were also made eligible to register as part of a public-private partnership (PPP) in sub-counties either in which there were no participating public secondary schools, where those government schools were over-crowded, or where pupils must travel very long distances to reach the closest government school. This policy is borne out in the 2013 EMIS data – 91 percent of sub-counties with no free USE government school have a PPP school, compared to only 52 percent of sub-counties that do have a free USE government school. To qualify, schools must be registered, certified, charge low fees (defined as 75,000 UGX ($22 USD) or less), and meet a set of criteria including having adequate infrastructure, a board of governors, and sufficient qualified teaching staff. Partnering private schools also become eligible to receive other support from the government including the provision of textbooks and other teaching materials.

For students to be eligible, they must have a score of 28 or better in their PLE exam, corresponding to an average passing grade in each of the four tested subjects. In our sample, six percent of students at government USE schools and eight percent of students at private USE (PPP) schools had actually failed to meet this threshold. The majority of students enrolled are funded through USE.

Government schools are entitled to 41,000 UGX ($12 USD) per term per student (in addition to other transfers to schools including teacher salaries), and PPP schools to 47,000 UGX ($14 USD) per term per student, on condition that they do not charge any other non-boarding fees. In practice, despite transfers from government and fees being prohibited for USE students, parents still report substantial fees paid to both government and PPP schools. Median reported annual household spending on school fees per child at secondary school was 360,000 UGX ($107 USD) for PPP schools and 150,000 UGX ($44 USD) at government schools (this includes registration fees and contributions to school development funds). Similar amounts are spent on books and uniforms in government and private schools

Table 1.1: Ugandan School Characteristics

School Type		Government	Elite Government	PPP	Private	International PPP	Total
Total Pupils (million)		0.56	0.12	0.46	0.5	0.01	1.65
Schools	All	947	121	825	1,698	23	3,614
	Sample	75	7	55	56	19	212
Average Pupils	All	592	1,002	554	294	488	456
	Sample	610	495	506	281	491	482
Fee Rate	All	190	975	96	260	209	228
	Sample	194	943	91	289	175	215
Pass Rate	All	0.37	0.78	0.38	0.57	0.49	0.47
	Sample	0.38	0.74	0.38	0.46	0.51	0.43
Repetition Rate	All	0.01	0.01	0.01	0.01	0.01	0.01
	Sample	0.01	0.04	0.01	0.02	0.01	0.02
Government Teachers	All	19	40	0	0	0	6
	Sample	17	27	0	0	0	7
Private Teachers	All	9	16	22	21	20	18
	Sample	10	14	21	21	19	17

Notes: All data is from the 2016 EMIS. Fee rate is in '000 Ugandan Shillings. Pass rate is the proportion of UCE candidates who achieve a Division I-III

Rollout of the USE programme amongst private schools was randomised, allowing for a high-quality estimate of the impact on private schools of accessing this public funding. Private schools that obtained public USE funding experienced greater enrolment growth, improved student performance (on low-stakes tests), but also more selection of better performing students at entrance. Despite the official eligibility requirements, there was no effect of USE registration on school governance arrangements (Barrera-Osorio, Galbert, Habyarimana, & Sabarwal, 2016).

1.3 Data

I compile data from our own 2016 management survey, with school characteristics taken from the 2015 Ark School Survey, official test score data for 2015 from the Uganda Examinations Board (UNEB), and further contextual data on schools from the 2013 Education Management Information System (EMIS), and the national population census in 2014 and 2002. Table 1.2 shows a summary of test scores and management scores by school type.

Management Survey

I measure school management quality using an adapted version of the Bloom et al (2015) and Lemos & Scur (2016) school management surveys[5]. Open-ended interviews are carried out, with responses scored against a descriptive rubric on a one - five scale for 20 question topics. These topics are grouped into four main

---

[5] Full details of changes made to the instrument are included in Appendix Al.3.

areas; target-setting, monitoring, operations (planning and leading teaching), and people (teacher) management.

- Operations (planning and leading teaching): this covers the leadership of teaching in a school, the use of differentiated teaching for a range of students, how schools use data and assessment to guide practice, and how education best-practices are adapted;
- Monitoring: this includes how the school tracks and monitors performance; whether there are systems and processes in place to identify and fix problems; and how stakeholders are involved in on-going quality improvement (students, teachers, community);
- Target setting: this includes how school targets are linked to student outcomes; specific targets for departments and teachers, how appropriate the targets are;
- People: how teachers are recruited, managed, supported and retained.

Table 1.2: School Characteristics (Survey Sample)

	Gov	Elite Gov	PPP	Private	IPPP	All	N
Number of Schools	82	7	62	48	24	223	223
Management Score	2.0	2.4	2.0	1.9	3.0	2.1	223
Operations	2.0	2.4	1.9	2.0	3.3	2.1	223
Monitoring	2.1	2.4	2.0	1.9	3.0	2.1	223
Target-Setting	1.7	2.4	1.8	1.6	2.7	1.8	223
People	2.1	2.2	2.1	2.1	3.1	2.2	223
Test Scores & Students							
Value-Added (z-score)	0	0.59	0.07	0.24	0.28		
Mean Test (UCE z-score)	0	1.21	-0.05	0.38	0.21		
Mean Students (2015)	563	552	427	281	510	458	223
SES Index (z-score)	-0.39	1.46	-0.07	0.26	-0.37	-0.09	210
Dropout Rate (S3-UCE, %)	0.20	0.22	0.23	0.20	0.34	0.21	205
School Characteristics							
Total Fees* (UGX)	72	109	79	114	107	88	223
% Religious	0.66	0.86	0.60	0.46	0.00	0.53	223
% Rural	0.83	0.57	0.55	0.48	1.00	0.69	223
Km to Kampala	211	180	193	167	192	193	223
% Heads with postgrad	0.39	1.00	0.18	0.13	.	0.25	223
Head Experience (Years)	10	19	9	7	.	9	197
Teacher Experience (Years)	7	10	6	6	.	7	199
Teacher Salary / Month ($)	107	-	-	101	98	-	55
Autonomy							
Admissions	0.71	0.86	0.89	1.00	1.00	0.86	211
Staff	0.69	0.50	0.98	0.98	1.00	0.86	223
Academic	0.45	0.53	0.57	0.63	0.64	0.54	208
All (Mean)	0.62	0.67	0.81	0.85	0.86	0.75	223
School Market							
Schools per capita**	1.4	2.4	1.7	1.9	0.8	1.6	214
School Age (years)	27	44	16	11	3	18	216
Distance to NTC (mean Km)	98	112	84	88	94	92	223
2002 Child Literacy Rate**	0.45	0.53	0.46	0.46	0.46	0.46	199

Notes: Management scores and school autonomy scores are from the WMS-style management survey. Test scores are normalised so government schools are equal to zero and other school types are shown as the unconditional difference in standard deviations. Test score data comes from UNEB for 2016 and 2015 and the Ark School Survey for 2014. School characteristics are from the Ark School Survey. School Market variables are from the EMIS, census, and Ark School Survey. Both final exam scores and value-added scores are first standardised at the individual student level by year, before taking means across all students in each school type. Data on teacher salary comes from EPRC 2016.

* Total fees comprise tuition fees plus fees for extra classes, uniforms, lunch, & 'other'.

** These school market variables are presented at the sub-county level

Table 1.3: Management quality by school type

	Schools	Management (Mean)	Management (SD)
School type			
Government (USE)	82	2.0	0.32
Elite Government (Not USE)	7	2.4	0.24
PPP (Private USE)	62	2.0	0.34
Private (Not USE)	48	1.9	0.32
International PPP (USE)	19	3.1	0.59
Location			
Kampala	4	2.24	0.31
Other Urban	22	2.10	0.34
Rural	173	1.95	0.33
Religious Orientation			
Not religious	95	2.2	0.62
Anglican/Protestant	61	1.9	0.32
Catholic	47	2.0	0.35
Other	11	1.9	0.34
Academic Selection			
No Selection	11	2.0	0.66
Academic Selection	204	2.1	0.48
Profit			
Not for Profit	163	2.1	0.52
For Profit	50	2.0	0.35
Headteacher qualifications			
Postgraduate	56	2.0	0.37
Graduate/ Bachelor's degree	142	2.0	0.32
Headteacher employment			
No other job	187	2.1	0.51
HT has 2nd job	28	1.8	0.22

Notes: This table shows average management quality score for different school types.

Each of the 20 scores depends on a series of individual questions that help build up an overall description of the concept being measured. This approach combines a rich open-ended discussion of management practices allowing for probing and clarification where necessary, with a quantitative framework to allow for comparison between schools. Scoring inevitably still depends on a subjective judgment by individual interviewers, and so substantial time needs to be taken in training enumerators, discussing in detail the level descriptors, and calibrating scores across interviewers across a range of practice interviews.[6]

The management survey was carried out with a stratified random sample of 199 schools from the 2015 Ark School Survey, plus the 19 international public-private partnership schools. The sample includes 82 regular government schools, seven elite government schools (not part of the free secondary education programme, high fee-charging, high socioeconomic status students), 62 public-private partnership (PPP) schools, and 48 fully private schools.

The survey was carried out in January 2016 by telephone from a call centre in Kampala, from a nationally representative sample of 305 schools (stratified by ownership and district), from which an overall response rate of 65 percent was obtained (199 schools). A list of school leader phone numbers were provided by the Ministry of Education. 29 percent of these numbers failed to connect or were not answered. Only six percent refused to participate in the survey. This response rate is substantially higher than that found in other countries, from a high of 58 percent in Brazil to just eight percent in the UK). A linear probability model (LPM), probit, and logit model all show that none of the main school

[6] Interviews were double-scored in training, with a correlation of above 0.9 between scores from different enumerators.

characteristics[7] from the first round survey are correlated with the probability of response for the second round management survey (Table A1.6).

Interviews lasted between 60 and 90 minutes. Around ten percent of interviews were double-scored by a research manager, with an average variation in double-marked overall scores of 0.1 - 0.2 points. Surveys benefit from being "double-blinded" in the sense that interviewers are not influenced by their physical impressions of the school or knowledge of school performance, and respondents were not aware of the rubric against which they were being graded. Telephone surveys have been demonstrated in other contexts to generate data that is statistically indistinguishable from in-person interviewing (Garlick et al. 2016, Bloom et al. 2012a).

We also asked a set of standard questions on school autonomy taken from the OECD PISA survey. Headteachers are asked who has the main responsibility for deciding on budget allocations, selecting teachers for hire, setting teacher salaries, deciding who to admit, which courses to offer, the content of courses, and which textbooks to use. Where the head teacher, school owner, or governing board are primarily responsible, this is coded as the school having autonomy over that area, whereas where the Ministry of Education is primarily responsible the school is coded as not autonomous. In line with our expectations, private schools and PPP schools have a similar level of autonomy, which is greater than the autonomy of regular government schools. On budget autonomy, almost all private schools and the majority of government schools claim to have school level autonomy. On salaries and hiring, almost all private schools report having autonomy, compared

[7] The characteristics tested are the number of students, average socioeconomic status of students, years of operation, location, average fees, head teacher experience and qualifications, teacher qualifications, and school type.

with 70 percent of government schools. Private schools are also more likely to report autonomy on admissions, course choice, and textbook choice. On course content only around a quarter of schools, whether public or private report having autonomy, with content most commonly being determined by the Ministry of Education.

Test Score Data

Students take national standardised tests at the end of primary (PLE) and then again at the end of junior secondary school (UCE). Prior to 2015 this data was not digitised and centrally stored. In 2015 the Ark School Survey visited schools and collected UCE scores directly from school paper records for a sample of schools in 2014 and 2013. In addition, the linked PLE score (from 2009 or 2010) for each student was obtained from school records. From 2015, the Uganda National Examinations Board (UNEB) provided a full national set of individual student UCE results, linked to their individual PLE result. This gives a total sample of 43,156 students across three years from 218 schools.

Students sit UCE exams in eight (or more) subjects. Their final classification is based on an average points score across their eight best subjects. Points are awarded based on the percentage mark in exams, with one point as the best possible score corresponding a mark of 80 - 100 percent on the exam, and nine points being the worst possible score, corresponding to a mark of 0 - 39 percent. Our main outcome variable is the aggregate point score across eight subjects (inverted so that positive coefficients mean a better result, and standardised so that the mean is zero and the standard deviation is one, to allow for easier interpretation of estimated coefficients).

PLE scores are scored in a similar manner from one to nine points for each of the four individual subjects (English, Maths, Science, and Social Studies). For the prior test score variable, again, I take the aggregate points score across these four subjects, invert it and standardise it.

Other School Characteristics

I also make use of a range of other school characteristics taken from the 2015 Ark school survey, including a school-average student socioeconomic status index (following Filmer and Pritchett, 2001), the total number of students, average tuition fees charged, and number of years in operation.

The Ark School Survey was carried out in 2015 with a nationally representative sample stratified across Uganda's four regions and across school type (public and private). Ten districts were sampled from each of the Central, Western, and Eastern regions, and six from the less populated Northern region. For each district ten schools were randomly sampled, of which four were government schools and six private schools.

Public schools are on average larger than private schools, though PPP schools are closer in size to public schools as they receive a government subsidy per pupil place. Schools of all types report charging tuition (and other) fees, despite this not being officially permitted for government schools and PPP schools. Of the fully private schools, around half are non-profit. 95 percent of schools use academic selection criteria. The majority of schools (55 percent) are religious. The majority of government schools are in rural areas, with private schools and PPP schools more prevalent in urban and peri-urban areas.

Average socioeconomic status of students is estimated with a household asset survey administered to students in the fourth grade of secondary school (S4) following Filmer and Pritchett (2001). This data is not linked to individual test score results as those students had already left the school, but instead gives an estimate of school-average socioeconomic status. Students are private schools are 0.15 standard deviations higher than average socioeconomic status.

Headteachers and teachers have fewer years of experience in private schools than in government schools, and are less likely to have higher qualifications, in line with private schools in general paying lower salaries and having lower job security than in the public sector.

School Markets

In order to understand the factors that affect management quality, I assemble a series of additional contextual variables about the markets that schools operate in. First, I measure competition as the total number of schools (taken from the 2013 Education Management Information System (EMIS)) per capita (from the 2014 census) within a sub-county. There are 2,792 secondary schools nationally and 1,382 sub-counties, giving an average of two schools per subcounty. In our sample the median school is in a subcounty that has three schools in total. Second, school age in years is taken from our school survey. All government junior secondary school teachers must have at least a qualification from one of these colleges or a university. Third, I use two measures of the local supply of skills, the distance from each school to a National Teacher College (NTC) is calculated based on the shortest distance between their GPS coordinates, and the local child literacy and enrolment rates are calculated from self-reports of literacy from the 2002 census for all children aged five to 18 years.

1.4 Descriptive Statistics

There are a total of 1,068 government and 2,546 private secondary schools in Uganda. Government schools have on average nearly twice as many students as private schools. Table 1.1 shows summary statistics for all schools (based on 2016 administrative data), and Table 1.2 shows more detailed summary statistics for schools in my survey sample.

On average schools in the nationally representative sample of Ugandan secondary schools score 2.0 points on the one to five scale, placing them above India and slightly below Brazil. This comparison should be taken with some caution due to the slight adaptations to the survey instrument. In other contexts schools tend to perform worse than manufacturing and retail firms. The best performing country that has been measured for school management is the UK at 2.9 (Bloom et al 2015). The distribution of schools in Uganda is roughly symmetrical, with very few schools in Uganda scoring above three points, which is similar to the distribution in India, but notably different to that in Brazil where despite low average management scores, there is an upper tail of high performance.

Management scores do not vary systematically for government, private, and PPP schools for either the aggregate score or any of the sub-components. Elite government schools (those not in the USE scheme) do score 0.4 points higher, and more dramatically a chain of internationally-owned PPP schools score one point (two standard deviations) higher.[8] The difference in overall management quality between elite government schools and others is present in their operations management (teaching quality control), target setting, and monitoring, but they are not better than average on teacher management. There is also substantial variation in management quality within school types. This variation is greatest for the international PPP schools, which is possibly explained by them being substantially newer than other schools (average of three years old). Table 1.3 presents average management scores for other school characteristics.

1.5 Empirical Approach

1.5.1 Management and Student Performance

In order to look at the relationship between school management and student performance, I estimate a student learning production function following (Todd & Wolpin, 2003), in which student achievement T is conceived of as a function of their ability A, and all present and past family inputs F, and school inputs S.

$$\mathrm{T} = \mathrm{f}(\mathrm{A} + \mathrm{F} + \mathrm{S}) \tag{1.1}$$

Management quality enters this framework as a school input as a form of intangible capital that affects the productivity of labour and capital, can be invested in, and can depreciate. Equation (1.1) can then be re-written such that

[8] This finding is supported by a separate study into the same school chain (EPRC 2016) that found substantially greater evidence of schools having a vision and providing performance reviews and feedback to teachers, in the international PPP schools than in domestic PPP schools (Table 13).

the partial derivative of test scores with respect to school characteristics is a function of management capital M, non-management labour L, capital K, and an efficiency term α.

$$dT/dS = f(\alpha, L, K, M) \tag{1.2}$$

In practice, estimation of (1.1) is impeded by the lack of measures of student ability and the full history of family and school inputs. A common solution is the estimation of a lagged dependent variable, dynamic OLS 'value-added' specification, in which a student's prior test score serves as a composite proxy variable for both their unobserved ability and all observed and unobserved prior home and school inputs, which allows for the estimation of the marginal effects of contemporaneous inputs.

Here then test score T of student i at school s at time t is related to their own lagged performance, student characteristics X_i, school characteristics S_s, and school management quality M_s. Some of these school characteristics (specifically average socioeconomic status of students and school fees) proxy both for family inputs and school inputs. I assume that management quality is persistent and unchanging across the four years for which there is test score data.[9]

$$T_{ist} = \alpha + \beta^1 T_{ist-1} + \beta^2 M_s + \beta^3 X_i + \beta^4 S_s + u_{ist} \tag{1.3}$$

In principle these estimates may be biased due to non-random sorting of students to schools and unobserved student heterogeneity that may be correlated with

[9] In practice there is annual turnover of headteachers of around 14 percent, and so I also estimate the model with a single year of test score data, finding similar results.

both dependent and independent variables. As I only have measurements for each student from two time points I am unable to estimate models that include both student fixed effects and a dynamic component controlling directly for prior performance. In practice however the size of this bias has been demonstrated to be small. Using simulated data (Guarino, Reckase, Stacy, & Wooldridge, 2015) demonstrate that 'naïve' dynamic OLS models are more robust than other more complex non-experimental estimators in recovering relatively accurate teacher effects. Using real data various studies have shown that simple value-added models can recover good approximations to experimentally identified parameters. Several studies compare lottery estimates of school effects with observational value-added estimates using the same data, finding very similar results (Abdulkadiroğlu et al. 2011; Angrist, Pathak, and Walters 2013; Deming 2014). Focused on teachers, several studies find that observational value-added estimates of teacher effects in one year of a study are unbiased predictors of experimentally obtained value-added estimates of teacher effects from a different year (in which students were assigned to teacher classrooms randomly in the second year) (Kane and Staiger 2008; Kane et al. 2013; (Buhl-Wiggers, Kerwin, Smith, & Thornton, 2016). Finally observational value-added estimates of the effect of private schools in Andhra Pradesh, India (Singh 2015), are very similar to experimental estimates from the same context and point in time (Muralidharan and Sundararaman 2015).

An important concern in the Ugandan context is whether there are differential rates of dropouts between better and worse managed schools. It may be that better managed schools are successful primarily at encouraging under-performing students to leave. I argue that this is unlikely – schools are typically funded on a per-pupil basis either directly through fees or through per-student government subsidy, giving them a strong incentive not to cut enrolment. Further, parents and the media judge schools primarily on the absolute number of top grades (Division One) achieved, and so schools are not penalised if they have a high number of low scoring candidates. I can also test this concern directly with the data on dropouts.

My results could also suffer from omitted variable bias due to missing data on other school inputs, such as the pupil-teacher ratio. For my main results to be unbiased, I require that the conditional correlation should be zero (or minimal) between omitted variables, the management quality index, and student outcomes. In addition, I argue that school fees are a good proxy for key inputs such as the number of teachers and indeed the pupil teacher ratio, particularly given the high proportion of school costs that are attributable to teacher pay.

1.5.2 Does Management Quality Explain the Private School Premium?

In order to explore the role that management plays in explaining the effect of different school types (government/private/public-private partnership), I follow the approach of Imai et al. (2010) within the framework of a Linear Structural Equation Model. Concretely, I test whether the effect of school type on learning is mediated by management quality. The direct effect of school type on learning is captured in β^4 in equation (1.3), in which I control for the effect of management. The indirect or 'mediation' effect of school type on learning is captured by the product of the coefficient of management on learning β^2 in equation (1.3), and the coefficient of school type on management quality, ε^1 in equation (1.4) below.

$$M_s = \alpha_2 + \varepsilon^1 S_s + v_{ist} \tag{1.4}$$

The total effect is then the sum of the direct effect and the indirect effect. Identification of the mediation effect relies upon the assumption that the correlation between the residuals of the two equations is zero. I then estimate a 'sensitivity parameter' ρ as the size of correlation that would be necessary for the true effect to be zero.

$$\rho = \text{corr}(u_{ist}, v_{ist}) \tag{1.5}$$

1.6 Results

1.6.1 Management and Student Performance

I find a clear positive correlation between school management and student performance. On average, a school with a one standard deviation higher management score is associated with a 0.07 standard deviation higher test score, after controlling for prior test scores, sex, and school characteristics (location, average student socioeconomic status, school size, and fees) (Table 1.4). These results are broadly in line with estimates from other countries (Bloom et al 2015).[10] This main result is small relative to the estimated effect of private schools in the context and some experimentally evaluated interventions in other contexts (P. Glewwe & Muralidharan, 2015), but importantly better management practice does not necessarily have a large financial cost, and varies substantially by school income level. The opportunity cost of teacher time may also be low, to the extent that they are not working at their full capacity. Surveys of teachers in Ugandan primary schools find that teachers are absent from the class during scheduled lessons 57 percent of the time (Bold et al., 2017). Experimental studies on management training suggest that interventions to improve the quality of school management can be highly cost effective (Fryer, 2017). Our sample includes a nationally representative sample of 199 schools, plus a 'top-up' of 11 schools from a network of international charity-run PPP schools. To show that the main result

---

[10] Estimates for other countries (Bloom et al 2015) are based on school-average test scores rather than individual student test scores, and so need to be shrunk to account for unobserved within-school variation in test scores in order to allow for a direct comparison with my student-level estimates. Collapsing individual student test scores to school-averages reduces the standard deviation across units by around half (based on 2012 PISA data). When scores are then standardised (to z-scores), 'effects' on school-average scores are therefore roughly twice as large as 'effects' on individual student scores. Adjusted estimates for within-school variation vary slightly across countries but remain close to 0.5 in both the PISA data and my Uganda data.

is not driven by the inclusion of these schools, I show in column 5 the main specification on only the locally run schools, finding little difference in the coefficient of management.

Table 1.4: Regression of Student Test Scores on School Management

	(1)	(2)	(3)	(4)	(5)
Management (Z-Score)	0.198***	0.085***	0.079***	0.070**	0.065**
	(0.067)	(0.027)	(0.027)	(0.030)	(0.029)
Year FE	Yes	Yes	Yes	Yes	Yes
Prior Test Score		Yes	Yes	Yes	Yes
Location Controls			Yes	Yes	Yes
School Controls			Yes	Yes	Yes
School Type				Yes	Yes
N (Students)	53,449	53,449	53,449	53,449	50,990
N (Schools)	210	210	210	210	199
N (Years)	3	3	3	3	3
R-squared	0.023	0.537	0.559	0.568	0.575

Notes: Column (5) excludes the international charity PPP schools from the sample, leaving a nationally representative sample of Ugandan secondary schools. Std. Err. adjusted for clustering by school. School controls include number of students, fee level & average student household asset index. Location controls include region fixed effects, and dummy variables for Kampala/Urban/Rural. Controls for headteacher characteristics are omitted due to missing data reducing the effective sample size, but including these controls does not substantially alter the results. * p<0.1, ** p<0.05, *** p<0.01

Breaking down the overall management index into the four main components suggests that target-setting and people management are each independently correlated with student performance (Table 1.5). These results are similar to those in the international study, where people management has the largest relationship with performance followed by target-setting, monitoring, and operations (though there is no a priori reason why the sub-components of management should have equal weight in systems with different binding constraints to improved performance). I also test the effect of a collection of five management sub-indicators highlighted by Dobbie and Fryer (2013) as components of success. This 'Dobbie-Fryer index' of sub-indicators is significantly correlated with performance after controlling for school characteristics, but with a smaller magnitude than the overall measure of management.

Table 1.5: Regression of Test Scores on Management sub-Components

	(1)	(2)	(3)	(4)	(5)
Management (Z-Score)	0.070**				
	(0.030)				
Operations (Z-Score)		0.043			
		(0.030)			
Targets (Z-Score)			0.076***		
			(0.028)		
Monitoring (Z-Score)				0.032	
				(0.023)	
People (Z-Score)					0.052**
					(0.025)
School Controls	Yes	Yes	Yes	Yes	Yes
N (Students)	53,449	53,449	53,449	53,449	53,449
N (Schools)	210	210	210	210	210
N (Years)	3	3	3	3	3
R-squared	0.568	0.567	0.568	0.566	0.567

Notes: Std. Err. adjusted for 210 school clusters. The overall management index is the mean of the four subcomponents, each separately standardised and entered independently here. School controls include number of students, fee rates, ownership, student household asset index, location, and noise controls for the survey enumerator * $p<0.1$, ** $p<0.05$, *** $p<0.01$

I test for a range of interactions of management quality with student and school characteristics, finding little evidence for heterogeneous effects, despite the prior belief that better managed schools might put more attention on high potential students, as school reputation tends to depend on the number of candidates achieving the top grade. Looking at interactions of management with school characteristics; ownership, location, size, average socioeconomic status, and level of fees are all statistically insignificant. The one statistically significant interaction is with the dropout rate between S3 and the final S4 exam, implicitly a selection effect rather than a treatment effect, which is greater in better-managed schools. One explanation for this could be that better managed schools might encourage students not to take the final exam if they are not expected to do well, even conditional on their prior performance. The effect of management remains of similar magnitude when excluding specific types of schools (such as elite government and international PPP schools).

Robustness

An obvious concern is that better managed schools may perform better for reasons besides management. Controls for student prior test score, student socioeconomic status, and school fees all reduce but do not eliminate the coefficient on management, suggesting that there is some selection bias in the effect of management on test scores before controlling for student intake. In the value-added specification applied here I assume that prior test scores account for unobserved student ability, as well as all past inputs, both home and school. As these prior test scores come four years before the final test score, I need a measure of home inputs for the four years between the two tests. Here I assume that the measure of student socioeconomic status and average school fees paid are able to serve as a proxy for home inputs. I do not have any measurements at the classroom or teacher-level, though the hypothesised effect of management on performance should work through improved teaching at the classroom level through better support and accountability for classroom teachers.

Selection on unobservables

In Table 1.6 below I implement the Altonji et al. (2005) / Oster (2016) bounding exercise, which estimates the amount of selection δ on unobservables that would be necessary for the estimated coefficient of management $\hat{\beta}$ on student test scores to be zero. The selection parameter δ is expressed relative to (as a percentage of) the degree of selection on observables. A selection parameter δ of 1 is suggested as a heuristic cut-off point – so we assume that selection on unobservables is likely to be not greater than selection on observables, given that covariates are typically selected purposively in order to account for as much of the variation in the dependent variable as possible. In this case this assumption seems reasonable, as a student's lagged test score alone accounts for more than 50 percent of the variation in test scores. The bottom two rows indicate that if we thought that achieving an R^2 of 1 was realistic, then selection on unobservables would only have to be 57 percent of the amount of selection on observables for a $\hat{\beta}$ of zero to be possible. However Oster (2016) argues that an R^2 of 1 is unrealistic and too demanding a hurdle (only 10 percent of results published in top four journals over the previous five years pass this hurdle), and a more achievable benchmark for R_{max} is 1.3 times the R-squared achieved in the most complete specification $\tilde{R}$. In this case, my result 'passes' this test, in that selection on unobservables would have to be greater (1.46x) than selection on observables for the coefficient on management to actually be zero.

Table 1.6: Altonji et al (2005) / Oster (2016) sensitivity analysis

Model with Controls	
$\tilde{\beta}$	0.070
$\tilde{R}$	0.568
Model without controls	
$\dot{\beta}$	0.198
$\dot{R}$	0.023
Sensitivity Parameters:	
$\delta(R_{max}= 1,\ \hat{\beta}=0)$	0.69
$\delta(R_{max}=\tilde{R}\times 1.3,\ \hat{\beta}=0)$	1.75

Notes: The sensitivity parameter δ is estimated as a function of the coefficient of management on student value-added β and the r-squared R in two models; with and without the full set of control variables.

Dropouts

An important concern in this context is whether there are differential rates of dropout between the two examinations for better and worse managed schools. The value-added specification will produce consistent estimates only if dropouts are caused only by time-invariant student characteristics. The overall decrease in the size of the cohort that started S1 in 2011 and entered S4 in 2014 is 16 percent[11]. First, I argue that the major causes of student dropouts are student-specific demand-side factors rather than being related to school quality. Of people who completed one of the first three grades of secondary school but did not take the UCE exam, 69 percent reported that they left school due to trouble paying fees. Just one percent reported leaving due to poor academic progress (Uganda National Panel Survey Wave 3, 2012).

[11] Analysis of the Uganda Bureau of Statistics 2015 Statistical Abstract

Common approaches to dealing with bias caused by attrition include Heckman selection models and inverse probability weighting of observations, which can produce unbiased estimates if 'selection' or attrition is caused by observable individual characteristics. As this student-level sample only includes those who have taken the UCE exam, I do not have data on students that did drop out, so cannot estimate the probability of attrition within the sample. As an approximation however, I can look at the national distribution of primary school leaving exam (PLE) scores by gender, and estimate the probability of individual dropout based on the relative proportions of each score by sex for the pre-secondary entry PLE results and the PLE results of those taking the secondary certificate. As I do not have a credible instrument (a variable that causes selection but not the outcome) I do not estimate the Heckman selection model, but instead apply inverse probability weighting. Relying on the full distribution of PLE scores rather than the distribution of PLE scores for students that have already started secondary school relies upon the assumption that this distribution is not substantially different. I argue that this is a reasonable assumption, as the PLE is optional and costly, and is typically taken only by students who do intend to progress to secondary school, for which it is a requirement. Weighting observations by their inverse probability does not substantially affect the coefficient on management.

A final check is looking at the correlation between the reported number of dropouts between S3 and the final S4 exam at the school-level, for which I do have data, and the school management score. There is no systematic relationship between this school-level measure of dropouts and school management. In my sample, this rate of dropout between students in S3 and those taking the UCE exam at the end of S4 is 21 percent, above the overall national rate of reported dropout from students enrolled in S1 in 2011 to those enrolled in S4 in 2014 was 16 percent, down from a higher dropout rate in previous years.

Finally, I note that schools are typically funded on a per-pupil basis (either directly through fees or through per-student government subsidy) and there is weak accountability for schools if students do perform poorly on final exams, giving them an incentive to discourage dropouts and keep students enrolled.

Test Score Measurement

Another concern here is the measurement of the dependent variable (UCE test scores), and whether any flaws in official test results as proxies for student learning is correlated with any of the independent variables. If a better managed school was only better at preparing students for exams without them actually learning any more, results for the effect of management on performance would be biased upwards. One check available for this is a question asked of Head Teachers about the amount of exam preparation carried out in schools. Controlling for exam preparation makes no difference to the coefficient of management on performance (Table A1.11). Any 'classical' measurement error in prior test scores will lead just to attenuation of the effect of these prior test scores on secondary scores. I also test alternative scaling of the test score measure. Using an ordinal logit across test grades produces similar results to the linear approximation used in the main specifications (Table A1.11).

Management Index

The main management index I use is a weighted average of the 20 sub-areas of management, first taking the average of sub-areas for each of the four main sub-components, and then taking the average of these four sub-components. The relationship between management and student performance is robust to aggregating the individual question areas of management in different ways, either by simple averaging across all 20, or by principal components analysis (Table A1.11).

Headteacher turnover

Here I have assumed that management quality is time invariant, but in practice there is some degree of headteacher turnover. I do not have data on the length of time that headteachers in my sample have served in their current school, but overall the turnover rate in government schools is 14 percent. Restricting my sample to test score data only from the same year that the management score was collected (2016) and discarding the test score data for the other years does not substantially alter the main results (Table A1.11).

1.6.2 Does Management Explain the effect of Private Schools?

Without controlling for management quality, private schools perform 0.28 standard deviations better than government schools, and PPP schools 0.14 standard deviations better. A common explanation for any better performance of private schools relative to government schools is better management. Here however I find no evidence that private schools or PPP schools are any better managed than government schools. Although school management varies substantially, there are few differences on average between major types of schools (unlike in OECD countries where ‘autonomous government schools’ - here referred to as PPP schools - score highest on management). In Uganda there are two exceptions; first a small number of selective elite government schools with high fees and wealthy students, that are on average 0.4 points better managed than other government schools, and second a chain of internationally-owned non-profit PPP schools run by the UK charity PEAS, which score 1.1 points better than average.

Elite government schools are substantially better resourced than average, which might explain their advantage (despite this holding after controlling for student SES and school fees). International PPP schools on the other hand have primarily the same level of resources as local PPP schools. One plausible explanation for this better performance is the notion of technology transfer from the international owners of the chain from the UK to Uganda (in line with findings that subsidiary manufacturing firms of multinational companies perform better than domestically owned firms (Bloom et al, 2014 and Bloom et al., 2012b).

This is likely supported by the existence of an effective within-network accountability system, based on a rigorous modern inspections regime that combines official examinations data (Hanushek and Raymond 2005; Hanushek et al. 2013) with subjective performance assessment (Hussain, 2015). Anecdotally, the supervision model for the international school chain includes detailed targets for a range of performance indicators, high-stakes accountability for head teachers with the removal of those under-performing and promotion of those successful, and on-going support and challenge throughout the year. Unfortunately I do not however have the necessary variation in this study to test this hypothesis.

To test whether the quality of management explains the difference in performance of school types, I run two regressions of test scores on school type and the full set of control variables, first without and then with the management quality variable (Table 1.7). Adding the control for management quality (Column 3) makes little difference to the coefficient on private or PPP schools, but dramatically reduces the size of the coefficient on international PPP schools, which loses statistical significance. Following the approach to causal mediation outlined by (Imai et al., 2010), I subject this latter finding to a sensitivity analysis. Although this framework confirms the finding that management mediates the entire effect of international PPP schools, a sensitivity analysis shows that the effect is not robust to substantial correlation between the error terms from the test score and the management regression. The threshold value ρ at which the mediation effect would be zero is just 0.07 (Table 1.8).

Table 1.7: Regression of Test Scores on School Type

	(1)	(2)	(3)
Elite Government (not USE)	1.413***	0.514***	0.482***
	(0.211)	(0.143)	(0.146)
Private (not USE)	0.483***	0.276***	0.278***
	(0.130)	(0.061)	(0.058)
Private PPP (USE)	-0.031	0.144***	0.150***
	(0.110)	(0.042)	(0.042)
International PPP (PEAS)	0.204	0.256**	0.108
	(0.160)	(0.116)	(0.126)
Management (Z-Score)			0.070**
			(0.030)
Year FE	Yes	Yes	Yes
Prior Test Score		Yes	Yes
Location Controls		Yes	Yes
School Controls		Yes	Yes
N (Students)	53,449	53,449	53,449
N (Schools)	210	210	210
N (Years)	3	3	3
R-squared	0.115	0.566	0.568

Notes: Std. Err. Adjusted for 210 school clusters. School controls include number of students, fee level, & average student household asset index. Location controls include region fixed effects, and dummy variables for Kampala/Urban/Rural. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 1.8: Mediation Regression of Test Scores on International Schools

	(1)	(2)
ACME (Management)	0.161***	0.184***
	(0.034)	(0.037)
Direct Effect (IPPP)	-0.012	-0.094**
	(0.040)	(0.040)
Total Effect	0.150***	0.090***
	(0.213)	(0.023)
Controls		Yes
Obs. (Students)	51,519	51,519
Obs. (Schools)	210	210
% of Tot Effect mediated	1.073	2.044
H0: ACME=0	0.000	0.000
Threshold ρ at which ACME = 0	0.933	0.088

Notes: The 'Average Causal Mediation Effect' (ACME) is the product of the coefficient of management on test scores, and international PPP (IPPP) schools on management. The direct effect is the coefficient of IPPP schools directly on test scores. The total effect is the sum of the mediation effect and the direct effect. The last row reports the threshold value of the unobservable ρ correlation, above which the true ACME would be zero.

1.6.3 What Explains Management?

Finally, going beyond school type I look for other factors that might explain variation in management quality. Starting with the accountability framework laid out in the 2004 World Development Report, we can think of two possible routes of accountability for public service providers that might lead to improved school management – a) the long route of accountability from citizens through the state then down to service providers, or b) the short direct route through consumer or user pressure on providers.

With regards to top-down accountability we observe little variation across schools - students from all schools take the same common entrance and exit examinations. Government schools are subject to a very weak, process-focused inspections regime. One part of this relationship where we do observe variation is the degree of autonomy that schools are provided with, a common focus of studies on school performance. Here I do have measures of school autonomy and can test the correlation between this measure and school management.

With bottom-up accountability, the responsiveness of school management (and value-added performance) to parent/customer demand depends on how we view the school choice decision. If parents are seriously interested in quality and value-added, then we might think of competition as driving up standards. In this case the model outlined in Bloom et al (2015), in which management is a technology that affects the productivity of inputs (capital and labour), provides several intuitive predictions. Management increases performance, but also there is likely to be (i) a positive effect of competition on management, (ii) a positive effect of firm age on management, as the result of a survival/selection process in which poorly managed firms are more likely to go out of business and close (and therefore not reach old age), and (iii) that management is increasing in the local supply of skills (as the cost of hiring good teachers is reduced). An alternative theory is one in which parents care primarily about school reputation. When schools are also able to select their pupils (as they are in this context), competition can lead to segregation by ability, and no actual increase in school performance measured as value-added (MacLeod and Urquiola, 2015).

I test these predictions in the specification below, in which management M is estimated as a function of school characteristics S, headteacher characteristics HT, and community characteristics C (including the number of nearby schools per capita, the distance to a National Teacher Training College, and the quality of schooling in the sub-county 13 years ago).

$$M_s = \alpha + \beta^1 S_s + \beta^2 HT_h + \beta^3 C_c + u_{ist} \tag{1.6}$$

The only consistently significant school characteristics are school autonomy, the local skills index, and school type (elite government schools and PEAS international PPP schools are both better managed than average, even after controlling for other characteristics). There is no statistically significant correlation between the degree of local competition (measured by the number of schools per capita in the local sub-county), school age, headteacher experience or education, or school size or socioeconomic status of students (Table 1.9).

I had expected to find that private schools would score more highly than government schools at least on people or teacher management, due to the explanation for greater efficiency in the private sector in similar contexts so frequently being due to greater accountability for teachers. Looking at individual items with the overall people/teacher management score, private schools do in fact score better than public schools on hiring and recruitment, but no better in the other items (attracting talent, rewarding and promoting high performers, and dealing with poor performers).

Table 1.9: Regression of Management on School and Market Characteristics

	(1)	(2)	(3)	(4)
Schools per capita (Z-Score)	0.014	0.040	0.043*	0.035
	(0.032)	(0.027)	(0.025)	(0.025)
Autonomy Score	0.266**	0.220*	0.226*	0.226*
	(0.112)	(0.121)	(0.122)	(0.123)
School Age in Years (Z-Score)	-0.033	0.041	0.033	0.026
	(0.039)	(0.034)	(0.036)	(0.039)
Skills Index	0.061**	0.024	0.037*	0.034
	(0.026)	(0.018)	(0.021)	(0.022)
HT is a rookie	-0.144**	-0.023	-0.025	-0.041
	(0.073)	(0.065)	(0.066)	(0.067)
HT has postgrad	-0.047	-0.008	-0.004	0.000
	(0.067)	(0.065)	(0.065)	(0.066)
Government (USE)		0.000	0.000	0.000
		(.)	(.)	(.)
Elite Government (non-USE)		0.220**	0.229**	0.156
		(0.098)	(0.097)	(0.132)
Private PPP (USE)		0.011	0.013	-0.002
		(0.080)	(0.080)	(0.077)
Private (non-USE)		-0.080	-0.077	-0.109
		(0.078)	(0.080)	(0.089)
International PPP (PEAS)		1.059***	1.071***	1.025***
		(0.181)	(0.177)	(0.272)
Size (Students: z-score)				-0.010
				(0.032)
SES (Z-score)				0.049
				(0.051)
Region FE	No	No	Yes	Yes
N	169	169	169	163
r2	0.076	0.449	0.460	0.350

Notes: Schools per capita is defined within each sub-county. The autonomy score is the average of 7 dummy variables for whether the school has autonomy over admissions, budgets, hiring, salaries, content, courses, and textbooks. School age is taken from the 2015 Ark survey. The skills index is comprised of 3 variables – the distance from the school to one of the 7 National Teacher Colleges (NTC), the local (sub-county) literacy rate in 2002 and the local (sub-county) enrolment rate in 2002 (of children aged 5-18 at the time). School controls include number of pupils, socioeconomic status, region, and urban location. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

We are left with a puzzle – that despite better management practices improving school performance at little extra cost, most schools do not adopt them. Some clues are provided by the literature on technology adoption in developing countries, which identifies a number of possible constraints to adoption (Foster and Rosenzweig 2010; Jack 2011). The informational constraint seems particularly important in this context – it may simply be that most school leaders are not aware of what good modern management practices are, and how they can be applied in schools. One piece of evidence for this hypothesis is the very low correlation (0.145) between head teachers' self-assessment of the quality of management in their school with my measure. Neither do these self-assessments of school management correlate with student performance. Another possibly important constraint is on the supply-side – where there is little widespread provision by either market or state of management training in this context for school leaders.

1.6 Conclusion

This paper adds to a growing literature on the importance of the quality of management practices for school performance. I present the first internationally comparable measure of school management quality from sub-Saharan Africa, placing the management quality of Ugandan schools in international context. Management matters for school performance, measured by growth in individual student test scores (or "value-added"). Further, though there is some level of higher spending which can lead to better management (as demonstrated by the better performance of elite government schools), amongst non-elite schools there is little correlation between school fees or other school resources and management performance, showing that in principle better management can be a low-cost strategy for improving learning outcomes. School management is not significantly better in private or public-private partnership schools.

I find few variables that matter for explaining variation in school management. Autonomy may provide the opportunity for better management, but is not in itself sufficient. An international PPP chain does manage to achieve substantially better management quality and correspondingly improved student test scores, which I argue is due to a better top-down accountability and performance management system, though I do not have the necessary variation in the data to confirm this hypothesis. Future research could usefully address this question of how to improve school management at scale, and the role that performance management systems and school inspections can play.

Appendix A1.1: Figures

Figure A1.1: Map of Ugandan School Locations

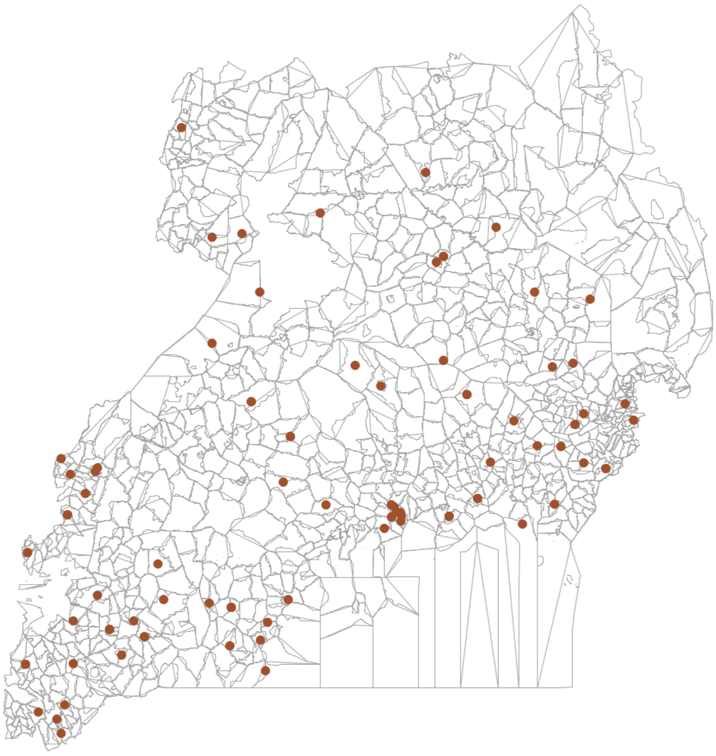
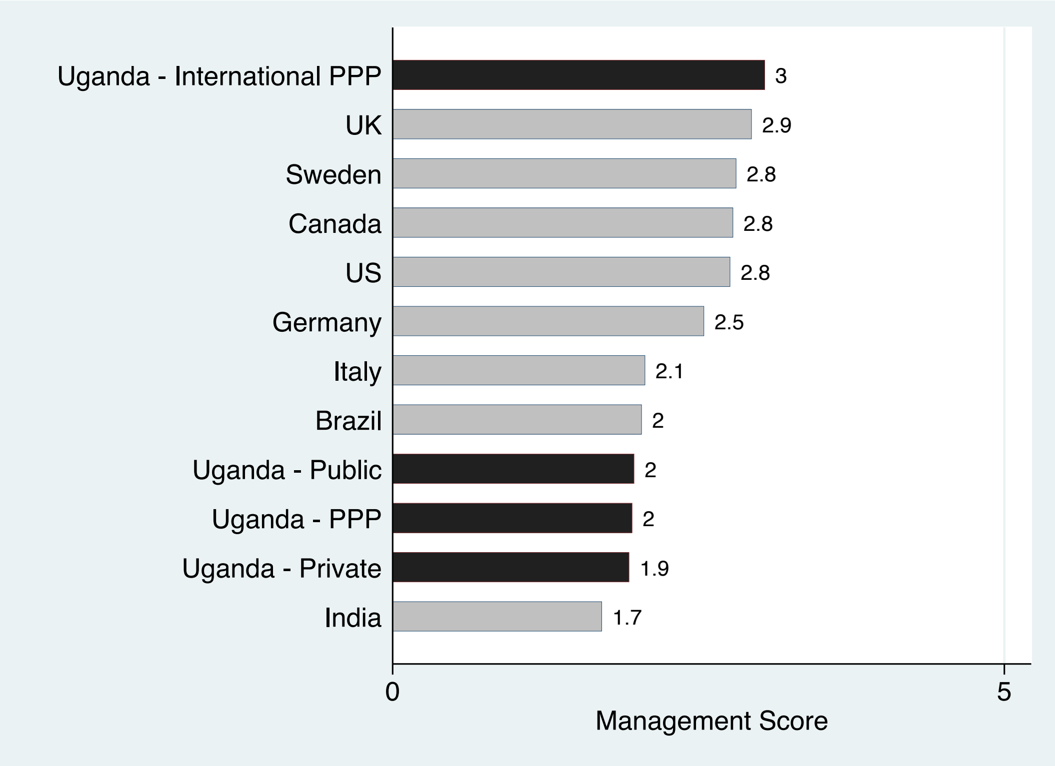
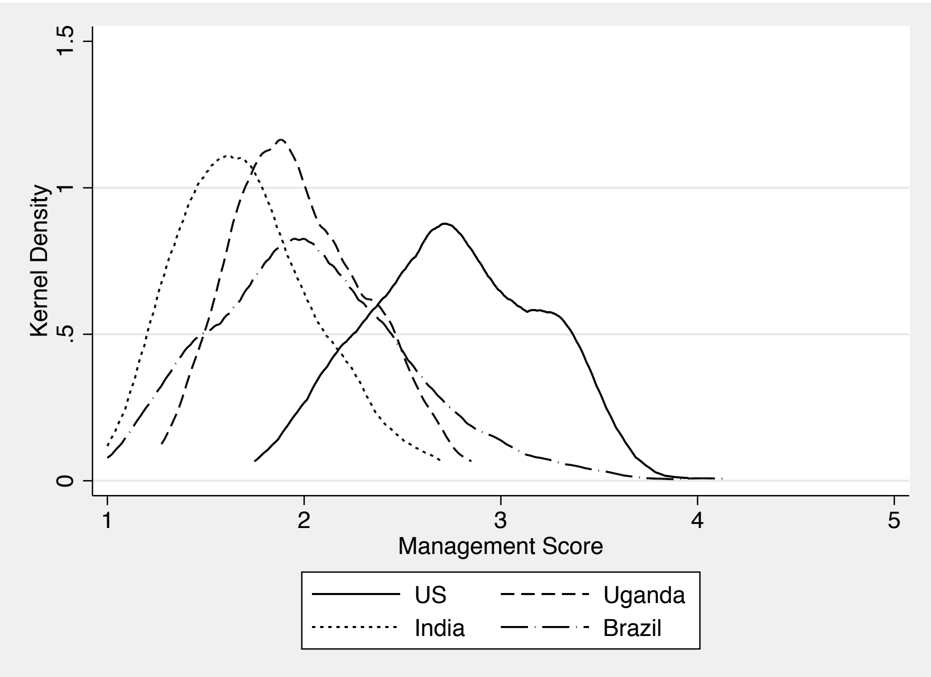


Figure A1.2: Management Score by Country and School Type



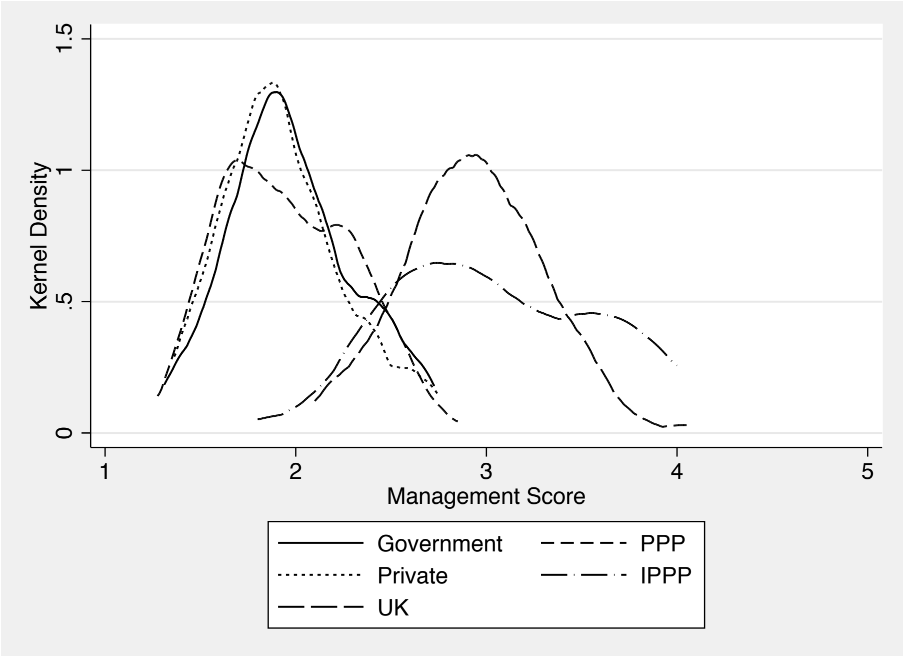
Notes: Scores from Uganda are from my survey, scores for other countries from Bloom et al (2015)

Figure A1.3: Distribution of School Management Scores within Countries



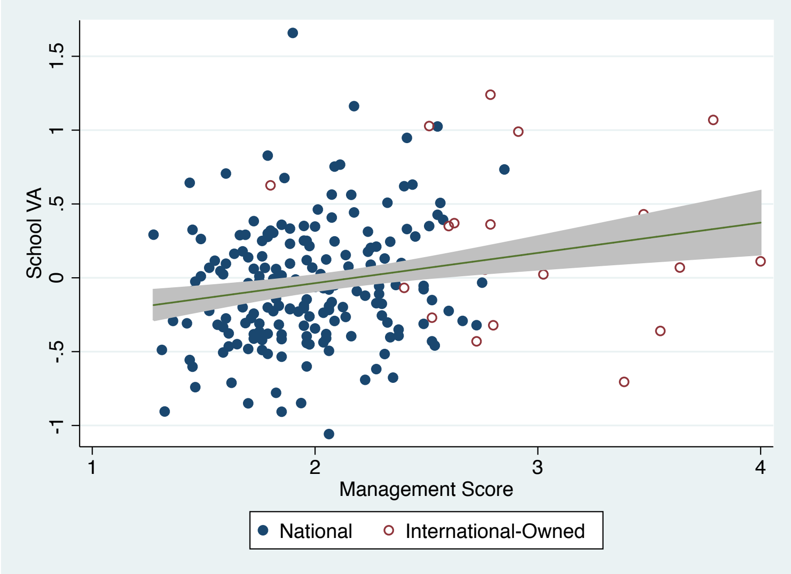
Notes: Scores for US, India, and Brazil are taken from Bloom et al (2015)

Figure A1.4: Distribution of School Management Scores by School Type



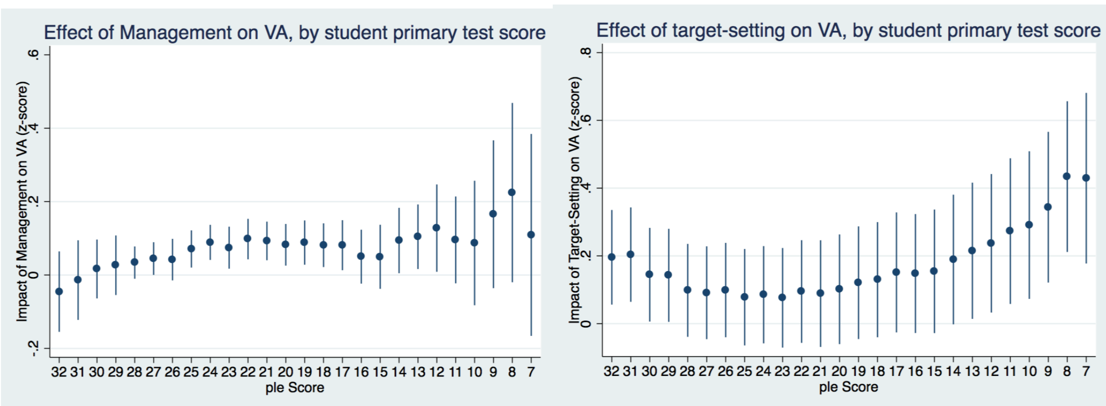
Notes: The distribution of management quality is presented here by school type. The distributions for public, private, and PPP schools all overlap, only elite government (omitted) and international PPP schools (IPPP) performing substantially better. Scores for UK schools are overlaid from Bloom et al.

Figure A1.5: Management and School Value-added



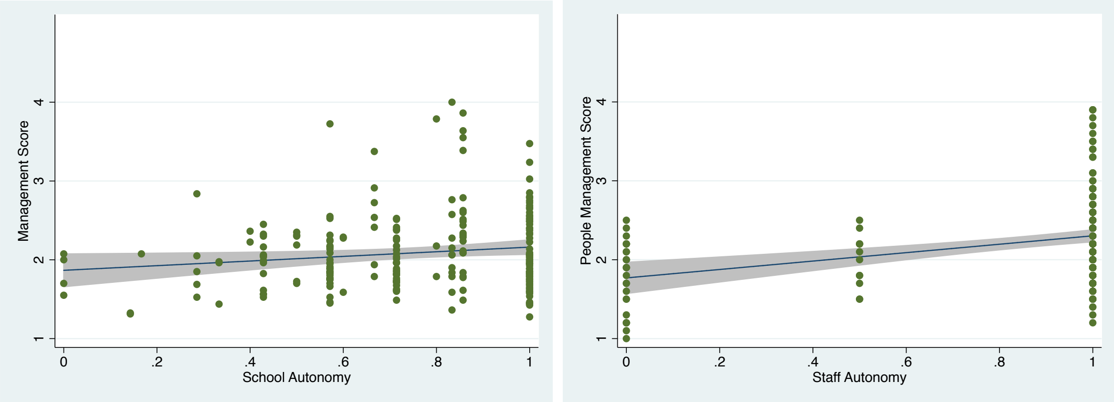
Notes: School VA is calculated as the simple school mean of residuals from a student growth regression, including controls for student prior test score, sex, and year.

Figure A1.6: Heterogeneous effects of management



Notes: Effect sizes estimated for each possible prior test score (PLE) with a piecewise regression, entering the interaction of each individual PLE score as a dummy variable multiplied by the school management score. Regressions control for school characteristics, with standard errors clustered.

Figure A1.7: Management Score and School Autonomy



Notes: Bivariate correlations between measures of autonomy and management show a weakly positive correlation for overall management (on the left) and a stronger correlation for staff/people autonomy and staff/people management (on the right).

Appendix A1.2: Tables

Table A1.1: Score by Individual Management Item

	Elite Gov (Non-USE)	Government (USE)	PPP (USE)	Private (Non-USE)	Foreign PPP (USE)	All Schools
Operations						
1.Planning	2.6	2.3	2.1	2.3	3.6	2.4
2. Leading teaching	2.2	2.0	1.9	2.1	3.4	2.1
3.Personalisation	2.5	1.8	1.8	1.8	3.2	2.0
4. Assessments & data	2.4	1.9	1.8	1.8	3.1	2.0
5. Adopting best practice	2.5	2.0	2.0	2.0	3.2	2.1
Monitoring						
6. Identifying problems	2.5	2.2	2.1	1.9	2.7	2.2
7. Performance tracking	2.4	2.0	1.9	1.9	3.4	2.1
Targets						
8. Target balance	2.6	1.8	1.9	1.8	3.0	2.0
9. Target Stretch	2.2	1.6	1.6	1.5	2.5	1.7
10.Accountability	2.3	1.7	1.8	1.7	2.7	1.9
People						
11. Hiring teachers	1.9	1.9	2.4	2.4	3.2	2.3
12. Attracting teachers	2.8	2.2	2.2	2.2	3.4	2.3
13. Rewarding teachers	2.1	2.2	2.3	2.1	2.6	2.3
14. Promoting teachers	2.1	1.9	1.7	1.9	3.1	2.0
15. Poor performers	2.3	2.2	2.0	2.1	3.3	2.3
Leadership & Ops						
16.Vision	2.6	2.2	2.0	2.2	3.0	2.2
17.Budgeting	2.6	2.4	2.2	2.0	3.0	2.3

Notes: This table shows average scores by school type for each of the individual sub-component questions that make up the overall aggregate management index.

Table A1.2: International Regressions of Student Test Scores on Management

	All (excl Ug)	Bra	Can	Ind	Swe	US	UK	Ug
	Score	Score	Score	Score	Score	Score	VA	VA
Mgmt	0.23***	0.10**	0.61	0.50**	0.24	0.17**	0.88**	
(z-score)	(0.044)	(0.050)	(0.368)	(0.243)	(0.206)	(0.080)	(0.369)	
School SD	0.49	0.59	0.47	0.49	0.54	0.51	0.50	0.46
Adj Effect Size	0.12***	0.06**	0.28	0.24**	0.13	0.08**	0.45**	0.07**
School controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pupil controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs-ervations	1002	472	77	152	82	133	78	210

Notes: Bloom et al (2015) estimate effects of management on school average performance (standardised to mean 0, standard deviation 1). In order to render these estimates comparable with the Uganda estimates using individual student data, I make an adjustment for the standard deviation of school-average test scores (calculated from 2012 PISA data for non-Uganda countries). The adjusted effect size is therefore an estimate from these studies for the effect of management on individual student performance.

Table A1.3: School Leadership & Management in PPP Schools

		None	Limited	Good	N
Evidence of school vision & mission	Foreign		4	7	11
	Domestic	6	5	5	17
Evidence of performance reviews & feedback	Foreign			11	11
	Domestic	3	6	8	17

Source: Economic Policy Research Centre (EPRC), 2016

Table A1.4: Regression of Individual Subject Test Scores on Management

	Eng	Mat	Che	Phy	Bio	His	Geo	Hum
Management (Z-Score)	0.069***	0.031	0.010	0.026	0.028	0.029	0.034	0.034
	(0.023)	(0.024)	(0.029)	(0.026)	(0.024)	(0.038)	(0.033)	(0.032)
School Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	34,437	34,416	34,400	34,352	34,361	18,497	18,647	18,103
N_clust	210	210	210	210	210	191	191	191
Dep var Mean	7.2	6.7	8.2	8.2	7.8	6.7	6.8	6.8
Dep var St Dev	1.7	1.8	1.3	1.3	1.4	2.2	1.7	1.7

Std. Err. Adjusted for 210 school clusters. School controls include size, fees, & student household asset index. Location controls include region fixed effects and dummy variables for Kampala/Urban/Rural.
* p<0.1, ** p<0.05, *** p<0.01

Table A1.5: Regression of Student Test Scores on the Dobbie-Fryer Index

	VA	VA	VA	VA
Dobbie-Fryer Index (Z-Score)	0.060***	0.041*		
	(0.022)	(0.025)		
Management (Z-Score)			0.085***	0.070**
			(0.027)	(0.030)
School Controls	No	Yes	No	Yes
N (Students)	53,449	53,449	53,449	53,449
N (Schools)	210	210	210	210
N (Year)	3	3	3	3

Notes: The “Dobbie-Fryer index” is my best approximation to the 5 key practices included in their actual index, taken from my school management survey. These include the sub-questions on data-driven teaching, the adoption of best practices, personalization of teaching, and leadership).
* p<0.1, ** p<0.05, *** p<0.01

Table A1.6: Selection regression of Choice to Participate in the Management Survey on School Characteristics

	OLS	Probit	Logit
Average UCE	-0.014*	-0.036*	-0.059*
	(0.008)	(0.021)	(0.035)
Average PLE	0.023	0.059	0.097
	(0.020)	(0.052)	(0.085)
Number of Students (2015)	-0.000	-0.000	-0.001
	(0.000)	(0.000)	(0.000)
Student HH Asset Index	-0.051	-0.132	-0.212
	(0.046)	(0.121)	(0.195)
Fees	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Headteacher Years Experience	0.001	0.002	0.003
	(0.004)	(0.011)	(0.018)
HT has postgrad	0.054	0.147	0.234
	(0.064)	(0.169)	(0.274)
Location Controls	Yes	Yes	Yes
N	323	323	323
r2	0.028		

Note: None of the location controls (dummies for region and urban/rural) are statistically significant in any of the specifications. * p<0.1, ** p<0.05, *** p<0.01

Table A1.7: Secondary School Enrolment and Dropout Rate

Year	S1	S2	S3	S4	S1 to S4 Dropout Rate	S3 to S4 Dropout Rate
2008	291,797					
2009	296,400	280,026				
2010	324,487	277,345	256,385			
2011	320,273	279,267	230,989	222,226	24%	13%
2012		296,297	259,003	216,754	27%	6%
2013			284,919	250,274	23%	3%
2014				268,253	16%	6%

Notes: Data from the Uganda Bureau of Statistics 2015 Statistical Abstract. Implied dropout rates are estimated by comparing the total size of each cohort as they progress through time and grades.

Table A1.8: Regression of Student Test Scores on School Management (Comparison of OLS and Random Effects Multi-level model)

	OLS (1)	OLS (2)	RE (3)	RE (4)
Management (Z-Score)	0.085***	0.059**	0.062***	0.059**
	(0.027)	(0.027)	(0.087)	(0.027)
Year FE	Yes	Yes		Yes
Prior Test Score	Yes	Yes	Yes	Yes
Location Controls		Yes		Yes
School Controls		Yes		Yes
School Type		Yes		Yes
N (Students)	41,818	41,818	43,156	41,818
N (Schools)	210	210	223	210
N (Years)	3	3	3	3
R-squared	0.529	0.569	0.529	0.569

Std. Err. adjusted for 210 school clusters. School controls include size, fees, & student socioeconomic status. Location controls include sub-region fixed effects and dummy variables for Kampala/Urban/Rural.

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A1.9: Heterogeneous effects of Management on Student Value-Added, by Student Characteristics

	(1)	(2)	(3)	(4)	(5)
Management (Z-Score)	0.066**	0.120*	0.051**		
	(0.030)	(0.071)	(0.024)		
Female x Mgmt	-0.014				
	(0.018)				
PLE x Mgmt		0.003			
		(0.002)			
PLE Division 1 x Mgmt			0.062		
			(0.046)		
Targets (Z-Score)				0.161***	0.057**
				(0.061)	(0.024)
PLE x Targets				0.008**	
				(0.004)	
PLE Division 1 x Targets					0.020
					(0.020)
School Controls	Yes	Yes	Yes	Yes	Yes
N (Students)	41,818	41,818	41,818	41,818	41,818
N (Schools)	210	210	210	210	210
N (Years)	3	3	3	3	3
R-squared	0.569	0.569	0.569	0.569	0.569

Std. Err. adjusted for 210 school clusters. PLE is a continuous variable with an aggregate points score ranging from 4 to 28. PLE Division 1 is a dummy variable for whether the student obtained the top grade in their primary exam. The positive coefficient on the interaction of PLE scores with Targets indicates that the effect of secondary school target-setting practice is greater for students with better predicted test scores (based on their primary test score). School controls include number of students, fee rates, ownership, student socioeconomic status, location, and 'noise controls' or enumerator * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table A1.10: Heterogeneous effects of Management on Student Value-Added, by School Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Management (Z-Score)	0.064**	0.062**	0.060**	0.067***	0.062**	0.070***
	(0.028)	(0.027)	(0.026)	(0.025)	(0.026)	(0.025)
Urban x Mgmt	-0.070					
	(0.073)					
Students x Mgmt		-0.010				
		(0.033)				
SES x Mgmt			0.045			
			(0.030)			
Dropout rate (Z-Score)				0.038		
				(0.025)		
Dropouts x Mgmt				0.080***		
				(0.028)		
Tuition Fees x Mgmt					0.023	
					(0.034)	
School Age x Mgmt						0.036
						(0.032)
School Controls	Yes	Yes	Yes	Yes	Yes	Yes
N (Students)	41,818	41,818	41,818	40,067	41,818	39,781
N (Schools)	210	210	210	205	210	203
N (Years)	3	3	3	3	3	3
R-squared	0.569	0.569	0.569	0.558	0.569	0.561

Std. Err. adjusted for 210 school clusters. Urban and Peri-Urban are dummy variables. School controls include number of students, fee rates, ownership, student socioeconomic status, location, and 'noise controls' or survey enumerator * p<0.1, ** p<0.05, *** p<0.01

Table A1.11: Robustness Checks

	(1) Main Spec	(2) Control for Exam Prep	(3) Ordered Logit	(4) Alternate PCA Index	(5) Alternate Un-weighted Index	(6) 2016 Only
Management (Preferred Index)	0.074**	0.077**	0.188**			0.079**
	(0.031)	(0.032)	(0.080)			(0.034)
% of Time on Exam Prep		0.000				
		(0.001)				
Management (PCA Index)				0.070**		
				(0.031)		
Management (Unweighted Mean)					0.054**	
					(0.026)	
Year FE	Yes	Yes		Yes	Yes	
Prior Test Score	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
School Controls	Yes	Yes	Yes	Yes	Yes	Yes
School Type	Yes	Yes	Yes	Yes	Yes	Yes
N (Students)	53,449	52,044	15,133	53,449	53,449	18,638
N (Schools)	210	204	208	210	210	190
N (Years)	3	3	3	3	3	1
R-squared	0.499	0.499	0.220	0.499	0.499	0.529

Notes: This table presents a series of robustness checks. Column 1 is the main preferred specification. Column 2 includes controls for the share of time spent on exam preparation. Column 3 uses one of four exam "Division" categories as an alternative outcome indicator, and an ordered logit specification. The reported R-squared is the pseudo R-squared. Column 4 presents results with the management sub-components aggregated using a Principle Components Analysis (PCA). Column 5 presents results with the un-weighted mean of the sub-components. Column 6 presents results for 2016 only. Std. Err. adjusted for clustering by school. School controls include number of students, fee level & average student household asset index. Location controls include region fixed effects, and dummy variables for Kampala/Urban/Rural. Controls for headteacher characteristics are omitted due to missing data reducing the effective sample size, but including these controls does not substantially alter the results. * p<0.1, ** p<0.05, *** p<0.01

Appendix A1.3: Edits to World Management Survey

The original World Management Survey schools instrument includes a rubric with level descriptors for one (worst), three, and five (best). The approach proposed by (Lemos & Scur, 2016) designed specifically for developing countries includes both a horizontal and vertical expansion of the tool, with level descriptors for half point levels at the bottom end of the scale (1, 1.5, 2, 2.5, 3, 4, 5) in order to capture variation in countries where scores are clustered at the lower end of scale, and including three separate sub-areas within each of the 20 question areas. During piloting for the Uganda survey it was decided to expand the original rubric to include level descriptors for each of the levels one to five, and to allow enumerators to score 0.5 points where they felt that responses fell between the two level descriptions, rather than describing explicitly what the 0.5 points were in the rubric. It was also opted to maintain the shorter set of 20 areas rather than expanding to 60, on the grounds that preventing respondent fatigue could outweigh any possible sacrifice in precision here. During pre-testing and piloting it was also decided to further simplify the original list of 20 areas to a combined and shortened list of 11 areas, to reduce excessive duplication and repetition of questioning and to limit the length of time required from a school head teacher. These changes are summarized in the table below.

Table A1.1: Changes to Original World Management Survey

Original WMS	**Adapted Instrument**	**Rationale for changes**
A. Operations		
Standardisation of Instructional Planning	Original category retained	Original categories retained. New category added to capture important school management role missing from original survey
Personalisation of Instruction and Learning		
Data-Driven Planning		
Adopting Educational Best Practices		
Instructional Leadership	New category	
B. Monitoring		
Continuous Improvement	Category retained	One category retained unchanged, the remaining four categories combined into one. In pre-testing we found that these questions/categories were very repetitive and overlapping and combined aspects of the categories into questions within a single category
Performance Tracking	Categories combined	
Performance Review		
Performance Dialogue		
Consequence Management		
C. Target Setting		
Target Balance	Categories combined	One category retained unchanged. Two categories combined where there is overlap. Some aspects of target interconnection were not relevant in this context – for example there are no district or national targets with which school targets could be interconnected. Two categories omitted - in pre-testing these questions were repetitive providing little new information
Target Interconnection		
Target Stretch	Original category retained	
Target Time Horizon	Categories omitted	
Target Clarity & Comparability		
D. People Management		
Recruitment	New category	Added from Lemos & Scur (2016)
Rewarding High Performers	Original category retained	
Fixing Poor Performers		
Promoting High Performers	Categories combined	Categories combined due to overlap and repetition in questions.
Continuing Professional Development		
Retaining High Performers		
Attracting High Performers	Category retained	

Chapter 2: Low Returns to Low-Cost Private Schools: Experimental Evidence from Delhi

A pre-analysis plan was filed with the American Economic Association's registry for randomized controlled trials, ID AEARCTR-0002239

2.1 Introduction

Enrollment in India's private schools rose by 16 million students between 2011 and 2015, surpassing half of all children in urban areas. Meanwhile enrollment in free government schools declined in absolute terms by 11 million. With median annual fees at around 10 percent of per capita income, parents often spend a large portion of their income on private schools. Providing poor families with access to these schools is a major focus of government policy under the 2009 Right to Education Act. While policy discussions often attribute demand for private schools to dissatisfaction with the quality of public education (PROBE Team, 1999; The Economist, 2015), other factors may be as important, including demand for geographic proximity or social exclusivity.

Prior experimental evidence from rural India has shown learning impacts from private-school vouchers were low or zero on average (Muralidharan & Sundararaman, 2015), but that returns increased with school density, leading the authors to speculate that in urban areas, where India's private school boom has been concentrated, "the effects of choice and competition may be considerably larger."

In this paper we exploit a lottery for private-school vouchers conducted in Delhi— precisely the type of dense urban schooling market where we might expect the biggest learning gains.[12] The voucher program, described in (Wolf, Egalite, & Dixon, 2015), offered vouchers to roughly 800 children in low-income neighborhoods in East Delhi which allowed them to attend a low-cost private school of their choice, tuition-free for five years. Take-up was reasonably high: winning the lottery led to a thirty-five percent increase in cumulative time spent in private school over the remaining five years of primary.

We find no evidence that access to private schooling generated gains in either cognitive or non-cognitive skills. Voucher winners saw no differential improvement in test scores on mathematics (point estimate of -0.04σ, standard error 0.05σ) or English reading (point estimate of 0.03σ, standard error 0.05σ), and suffered a negative impact on Hindi reading scores (point estimate of -0.09σ, standard error 0.05σ). Negative impacts on Hindi are somewhat unsurprising, as government schools are overwhelmingly Hindi-medium whereas 49 percent of private schools in our sample are English medium. It is notable, though, that private schooling led to no statistically significant gain in English to offset this loss in Hindi literacy.

Given these point estimates and standard errors, we interpret our null results as evidence of zero individual-level impact from school choice *per se*. However, two important caveats apply. First, winning a voucher increased cumulative exposure to private schooling by just over a third during the course of primary school. This implies that even if our intent-to-treat (ITT) estimates allow us to reject policy-

[12] Our study districts have 12 schools per km^2, compared to fewer than 1 school per km^2 in the Muralidharan and Sundararaman (2015) study districts.

relevant effect sizes from receiving a voucher of this magnitude, the confidence interval on the treatment-on-the-treated (TOT) effect of actually attending a private school is much wider. Second, while attrition was reasonably low at 13 percent overall, it was nearly 7 percentage-points lower among lottery winners. Thus the (Lee, 2009) bounds on our effects are quite wide, and leave open the possibility of larger effect sizes if, for instance, non-recipients of vouchers with particularly bad academic outcomes were significantly more likely to exit the sample.

We provide tests of three possible explanations for these results, finding some support for heterogeneous effects and asymmetric information. Overall, our experimental results on learning outcomes are broadly consistent with existing evidence in India that suggests that learning gains from private schooling are small and only apparent in some subjects and regions.[13] Analyzing longitudinal data on pupils in Andhra Pradesh, (Singh, 2015) finds some learning advantages in private schools (substantially higher learning gains in English and some modest and inconsistent advantages in Mathematics and Telugu), but consistent with our results finds negative effects on the local language (in his case Telugu, in ours

[13] Our results contrast somewhat with the original analysis of this voucher experiment by Wolf et al. (2015), which was based on two years of program exposure. Looking at intent-to-treat (ITT) effects, Wolf et al. (2015) report statistically significant learning gains of 0.15v on English, and insignificant gains of 0.07v in mathematics and 0.4v in Hindi, as well as larger and uniformly significant effects for the sub-sample of female students. In contrast, we find fairly precise zeros on all subjects, except for negative impacts on Hindi, and no notable gender differentials. We attribute this differential to three innovations in this paper. First, we test children after six years to study the effects of longer exposure to the program, thus results may naturally differ. Second, we invest in more intensive tracking of students, reducing the rate of attrition from 29 percent in Wolf et al. (2015) to 13 percent. Third, we base our selection of dependent variables and regression specifications on a pre-analysis plan registered prior to data collection in 2017 to preempt concerns about researcher degrees of freedom in selection of outcome variables and choice of sub-group analysis. We also supplement the primary academic outcomes with additional questions to provide further insight into the potential impacts of private schooling, including aspirations, psycho-social skills, well-being, and attitudes.

Hindi) in English medium private schools. As noted above, experimental evidence from the same state (Andhra Pradesh) presented in Muralidharan and Sundararaman (2015) shows no impact of private schooling on Telugu (the local language), Mathematics, or English, and positive impacts only on Hindi, a subject not taught in government schools at all.[14] Notably, while the experimental results from Muralidharan and Sundararaman (2015) are entirely rural, Singh (2015) finds no evidence of a private school advantage in urban Andhra Pradesh. Similarly, Desai et al. (2008) estimate positive impacts of private schooling on reading and mathematics overall using national survey data across India and an instrumental variable strategy, but no effect in Delhi specifically. Exploiting the Right To Education (RTE) lottery in the South Indian state of Karnataka, Damera (2017) finds no effect of attending private schools on learning, but a positive effect on student self-efficacy. Mis-targeting potentially mutes the results from Karnataka; although the RTE lotteries are intended for poor and marginalized groups, over ninety-percent of lottery losers pay out of pocket to attend private primary school.[15]

[14] Davies (2017) returns to the experiment described in Muralidharan and Sundararaman (2015) and finds that private school lottery winners in Andhra Pradesh were more likely to hold stronger market-oriented beliefs, be willing to pay more for private services, and express a preference for private service provision. We measure these outcomes only for parents, and find some corroboration of these findings.

[15] Effects may, of course, vary across contexts. In Punjab, Pakistan, Andrabi et al. (2011) estimate private school effects with value-added models, finding positive effects on learning. Looking further afield, a global meta-analysis considered 19 randomized trials spanning 11 voucher programs, including two studies from India already mentioned (Wolf et al., 2015; Muralidharan and Sundararaman, 2015), one from Colombia where vouchers for private secondary schools in an urban setting generated significant learning gains and labor market benefits (Angrist. et al., 2002, 2006), and eight from the US (Shakeel, Anderson, & Wolf, 2016). Formal meta-analysis found small positive effects overall, with larger effects outside of the United States and in programs with more years of treatment. A subsequent study from Louisiana not included in the Shakeel et al. (2016) review finds that participation in a school choice lottery lowered mathematics scores by 0.4 standard deviations (Abdulkadiroğlu, Pathak, & Walters, 2018).

Why do parents clamor for private schools if they generate little or no additional learning gains? We explore three hypotheses. First, that there may be non-learning benefits from private schools that motivate parents. Perhaps parents prioritize non-cognitive skills which are strongly associated with later-life outcomes (Deming, 2017), and which private schools instill. However we find no gains in a range of outcomes, including social awareness, self-management, aspirations, or growth mindset. Alternatively, parents may turn to private schools for non-educational attributes, such as access to a higher-status social network or physical amenities like textbooks and nicer classrooms, either as ends in themselves or signals of quality (Jacob et al., 2017; MacLeod and Urquiola, 2015).[16] We find no evidence that winning the lottery significantly changed pupils' social networks, as measured by the composition of their five closest friends. Treated students did however attend schools with slightly better amenities, including a reduction in class size of three pupils.

Rather than delivering higher quality—by whatever definition—the private sector may simply fill gaps in service provision where no public schooling is available (Barrera-Osorio et al., 2017). Consistent with this notion, evidence on private school enrollment decisions in rural Pakistan suggests distance matters enormously, with the probability of choosing a school falling more than 4.1 percentage points per 500 meter increase (Carneiro, Das, & Reis, 2016). By contrast, the price elasticity is relatively low: just -0.5 for girls and -0.2 for boys. If lack of state provision drives private school demand in Delhi, we would expect voucher winners to attend schools closer to their homes than other students.

[16] (Rao, 2018) shows that reservations for marginalized students in Delhi classrooms significantly reduce prejudice among their more elite peers, consistent with the contact hypothesis. Here we focus on the marginalized students themselves

Our data show no such effect, and distance to school is unaffected by lottery status.

Second, we present a simple conceptual framework that allows for richer and poorer parents to have different outside options. We might expect poorer children to go to low-quality, free government schools in the absence of vouchers. By contrast richer households might have otherwise gone to higher-cost and higher-quality private schools. We do find evidence of heterogeneity in treatment effects. Students predicted to have high spending on school have large negative treatment effects as they are induced to switch to lower-cost, lower-quality schools. Students predicted to have very low spending on school by contrast have positive treatment effects.

Third, an alternative explanation for our results is imperfect information. Parental demand may be based on limited or incorrect information about school quality. Ample evidence suggests school choices and student learning may be altered by provision of better information on relative school quality (Andrabi et al., 2017; Avitabile and de Hoyos, 2018), though results are not uniformly positive (Mizala & Urquiola, 2013). Our results reflect an environment where detailed information is lacking, and voucher-winning parents appear to change their minds about schools over time. Take-up of the vouchers was high in the first year, leading to a fifty percentage-point increase in the probability of attending a private school. But that gap declined by half by the fifth and final year of the program. A year after the program stopped, lottery winners were only five percentage-points more likely to enroll in private secondary schools. We also find evidence that initial school choice is more closely linked to raw test scores than value-added. Take-up is more likely to fall over time when households initially choose private schools with larger class sizes and fewer desks.

The following section describes the voucher lottery, designed to mimic policies mandated by—but often not enforced—under India's RTE Act. Section 2.3 describes our data, including the longitudinal tracking efforts over six years, the learning measurements, and other outcomes that we measured to explain sources of demand for private schooling. Section 2.4 presents the basic empirical specification and Section 2.5 reports our main results, starting with the basic intent-to-treat effects that are mostly zero for our primary outcomes, followed by pre-specified tests for heterogeneous effects and various robustness checks. Section 2.6 explores potential explanations for the null result.

2.2 The voucher lottery

The voucher program "Ensuring Access to Better Learning Experiences" (ENABLE), was administered by the UK-based charity Absolute Return for Kids (Ark) and the Centre for Civil Society. The experiment launched in the 2010-2011 school year and provided vouchers to winning students for five years until the 2015-2016 school year. The program operated in low-income neighborhoods in the East and North East districts of Delhi.

The ENABLE program was designed to mimic and hence test the model of the nationwide Right to Education Act 2009 Section 12(1)(c). This federal policy mandates all private schools to reserve 25 percent of their places for economically and socially weaker sections of society, with government providing a per child reimbursement equivalent to either government per child spending or the school's fees, whichever is lower. The Act was challenged by private schools in the High Court and only began implementation in 2012, and even then only in some states. 18 of 34 states and union territories have no schools participating. In Delhi, 48 percent of private schools were participating in 2013/14 (Sarin et al., 2015). Overall government currently finances 600,000 places at private schools through the program. Entry to the program is only in Grade 1 but each cohort will be funded up to Grade 8, so the expectation is that the total number of supported students will reach eight million in equilibrium.

A total of 1,776 eligible children applied for the program, of whom 1,618 were deemed eligible and ultimately entered the lottery. To qualify, students had to be aged between five and seven, from "economically weaker sections" (family income not exceeding Rs. 8,000 per month—about \$129 U.S. dollars), and not already enrolled in a recognized private school (they could be either attending a government school, an unrecognized private school, or not enrolled at all.) Overall children in our sample are from households below average income and wealth. Average monthly per capita income in our sample is Rs. 1,584, compared to a Delhi average of Rs. 2,654 (Table 2.1).

Table 2.1: Comparing experimental sample to Delhi as a whole

Variable	Sample (Endline)	Delhi
Scheduled Caste	0.21	0.27
TV	0.85	0.9
Fridge	0.53	0.76
Mother graduated Primary School	0.46	0.82
Average monthly income per capita	1,584	2,654
Annual household Spending Per Pupil	9,716	16,442

Note: Data for Delhi on Caste, television and fridge ownership, and mother's education are from the 2015/16 Demographic and Health Survey. Data for Delhi on average income (Rupees) are from the 2010 (66th) National Sample Survey. Data for Delhi on household education spending (Rupees) is from the 2014-15 (71st) National Sample Survey.

A random number generator sequentially selected students to be treated, stratified by ward (wards are electoral units with an average population of around 60,000 people. Students are drawn from 25 wards in East and North-East Delhi). If one sibling in a household won a voucher, all of their eligible siblings were also automatically awarded a voucher and withdrawn from the pool of unselected children, thus we cluster our errors at the household level. Students were selected sequentially until at least 50 percent of children in each ward were treated. This design meant that children with eligible siblings were slightly more likely to be in the treatment than the control group. In total 815 students were treated (50.4 percent of the sample).

Families of lottery winners were provided with a handbook listing all 105 schools and their address, principal, number of grades, average class size, medium of instruction, and toilet facilities. Parents were asked to rank their top five schools from the full list of 105 schools, with a second lottery held to assign students to schools if demand exceeded available places. In practice this was only necessary in a few cases, and the overwhelming majority (84 percent) of lottery winners were assigned to their first choice school.

A total of 105 private schools enrolled in the program from a total of 389 private schools in the 25 experiment wards. There are also 337 public schools in these wards. Schools committed to accept any and all voucher winners who applied to their school with no discrimination, admissions test, or additional fees. Each voucher was worth up to Rs. 7,300 (around $117 U.S. dollars), redeemable for annual tuition fees (up to Rs. 4,800), two sets of uniform (2 vouchers of Rs. 300 each), school textbooks (Rs. 900), and one meal per day (10 vouchers of Rs. 100

each). The fee level was set high enough that it would cover the fees at close to 75 percent of private schools in the experiment location. All private schools were then invited to participate. The voucher amounts increased by around 10 percent per year to account for inflation. The value of the initial voucher was around 15 percent of baseline household income. All participating schools charged fees at or below the Rs. 4,800 voucher level, and are in the low end of the fee distribution for private schools in Delhi. Average annual household spending on primary school in Delhi is more than twice the value of the voucher (Rs. 16,442 - MOSPI - NSSO, 2015). Average government spending per student in government schools is also twice as high as the ENABLE voucher (Rs. 14,615 - Pritchett and Aiyar, 2014). Vouchers were personalized for each student, produced by Edenred, an international voucher service provider operating in 100 countries, and followed Reserve Bank of India guidelines including a number of security features to reduce any risk of misuse (see Figure A2.2). A census of schools conducted in the same part of Delhi (North Shahdara) in 2004 found average teacher salaries were Rs. 10,072 per month in government schools, compared to Rs. 3,627 in recognised private unaided schools (Tooley & Dixon, 2007). A 2014 survey of private schools in which ENABLE voucher winners were enrolled was conducted. The survey covered 87 schools with treatment students and 35 schools with control students. The mean number of students in these schools was 175, of whom 42 percent were female. The median teacher had 4 years of experience. Schools reported an equal number of lessons per week in Hindi, English, and Mathematics, averaging 6 lessons of 35 minutes per week for each subject. Schools were allocated an average of 7 students each; by 2014 the number of voucher winners still at schools with voucher winning students had dropped to 5 per school. On entry this was around ten percent of the average grade one class of 68 students.

2.3 Data

We combine data from our endline survey, administered six years after the start of the lottery, with the baseline survey of all students collected for (Wolf et al., 2015), administrative data on initial school choices of treatment students, data on schools from a 2014 survey of treatment schools, and administrative data on schools from the District Information System for Education (DISE).

For our endline survey we administered three survey instruments: one for lottery students, one for their guardian, and one for their sibling. The student instrument tests English, Hindi, and Mathematics skills, alongside a set of non-cognitive measures, including growth mindset, self-management, social-awareness, and aspirations, as well as life satisfaction, self-perception, and social networks.
The English and Hindi assessments repeated items from the baseline survey allowing us to place the baseline and endline scores on the same scale and estimate absolute learning progress. At baseline, English and Hindi assessments were administered in an oral reading fluency format in which students simply read words aloud from a sequence. At endline, these same instruments were used along with a subset of publicly available items from the 2011 Progress in International Reading Literacy Study (PIRLS). One set of these items was translated into Hindi. The PIRLS questions are focused on testing reading comprehension by asking students questions about a short story, and are aimed at Grade Four students. In addition we include mathematics items from the Andhra Pradesh Randomized Experiment Study allowing us to compare our results with those from that voucher study (Muralidharan and Sundararaman, 2015). As (Tooley, 2016) notes, the Muralidharan and Sundararaman (2015) experiment

administered mathematics tests in the language of instruction, which varies systematically by school type. To minimize the risk of confounding school type and the language of testing, our mathematics instrument was presented side-by-side in both English and Hindi to all students. Figure A2.3 provides an example mathematics question from the exam. Questions on aspirations were drawn from the Young Lives Round 4 survey questionnaire - specifically, students were asked to "Imagine you had no barriers and could study for as long as you liked, or go back to school if you have already left. What level of formal education would you like to complete?". Parents were asked, "Ideally, how many total years of education would you like/expect your child to complete?". Items on other non-cognitive skills were drawn from the California CORE districts project (West, Buckley, Krachman, & Bookman, 2017), with four questions to measure each of the three concepts of grit, self-management, and social awareness. We measure life satisfaction using the Cantril Ladder, drawn from the Gallup World Poll, which we also adapted to ask about student satisfaction with their school. The sibling test was administered to the next youngest sibling who was present. In the case in which there were no younger siblings, the next oldest sibling was tested. Siblings were given the standard ASER Hindi and Mathematics assessments, which evaluate proficiency on a one to five scale. Siblings were also given an oral reading fluency test based on a list of English words.

Endline data was collected during household visits by teams of two enumerators in two phases. Enumerators took the following steps to locate children: first, all phone numbers on record were dialed; second, all existing addresses on record were visited; third, neighbors of the family's past addresses were asked for information for locating the family; and fourth, the school on record for the child was asked for information for locating the family. The first phase of endline data collection took place in May and June 2017, tracking 1,304 of 1,618 children. The second phase was conducted between September and October 2017, tracking an additional 101 children, for an overall attrition rate of 13.2 percent.

Baseline data was collected during household visits before the lottery was conducted by Wolf et al. (2015). Eligible students were given achievement tests in Hindi reading, English reading, and mathematics. Table 2.2 presents evidence on balance between treatment and control groups. Note we see statistically significant imbalance in age (with treatment pupils 1.5 months older on average) and on Hindi scores (where treatment pupils read 0.07 more words per minute). Inclusion or exclusion of baseline controls to adjust for this imbalance does not affect our main results.

We link the schools from the ENABLE experiment to administrative data from India on school characteristics from the District Information System (DISE). We are able to match 90 percent of the schools who selected into the voucher program, and 92 percent of the schools to which students were ultimately allocated.

Table 2.3 compares the private schools that were part of the ENABLE voucher program to other schools in the same village ID from the DISE data. Details on the variable construction can be found in Appendix A2.1. The DISE administrative data shows that schools enrolling voucher winners are on average smaller than non-participating private schools (253 students compared to 396). They have similar proportions of female and scheduled caste students, and similar levels of infrastructure and materials (books, water, toilets, playgrounds, and libraries). Private schools participating in the ENABLE program were, however, less likely to be English medium (49 percent) than other private schools (72 percent). Across Delhi almost all government schools are taught in Hindi (94 percent), and almost all private schools are English medium (86 percent) (ASER, 2014). Although most private schools are advertised as English medium, in practice most lessons are taught in both Hindi and English, with teachers having a relatively poor command of English and primarily teaching through Hindi translations of an English language textbook (Endow, 2018; Bhattacharya, 2013).

Table 2.2: Balance Test of Baseline Student Characteristics

Variable	(1) Control Mean/SE	(2) Treatment Mean/SE	T-test Difference (1)-(2)
Age (months)	65.781 (0.340)	67.291 (0.345)	-1.511***
Female	0.473 (0.018)	0.468 (0.017)	0.006
Baseline Monthly Income (1000s)	4.024 (0.073)	4.066 (0.082)	-0.042
Baseline Asset Index	0.031 (0.058)	-0.029 (0.057)	0.059
Hindi	-0.000 (0.037)	0.115 (0.042)	-0.115*
English	0.000 (0.038)	0.055 (0.040)	-0.055
Math	0.000 (0.037)	0.046 (0.043)	-0.046
N	784	834	
Clusters	689	649	
F-test of joint significance (F-stat)			1.659

Note: This table shows t-tests among baseline variables between treatment and control groups as well as an F-test for joint significance. Standard errors are clustered at the household level, and ward fixed effects are used.

Table 2.3: Selection of Participating Private Schools

Variable	(1) Voucher Schools Mean/SE	(2) Public Mean/SE	(3) Nearby Private Mean/SE	T-test Difference (1)-(2)	(1)-(3)	(2)-(3)
Total Students	252.656	592.344	396.010	-339.688***	-143.354***	196.334***
	(15.946)	(18.552)	(21.471)			
Portion Female	0.396	0.494	0.391	-0.098**	0.004	0.102***
	(0.006)	(0.024)	(0.005)			
Portion SC/OBC	0.092	0.207	0.063	-0.115***	0.029	0.144***
	(0.019)	(0.011)	(0.008)			
Portion Disabled	0.000	0.005	0.001	-0.004***	-0.000	0.004***
	(0.000)	(0.001)	(0.000)			
Portion Repeaters	0.030	0.036	0.003	-0.006	0.027**	0.032***
	(0.017)	(0.002)	(0.002)			
Co-Ed	1.000	0.234	0.990	0.766***	0.010	-0.755***
	(0.000)	(0.023)	(0.006)			
English Medium	0.490	0.033	0.717	0.457***	-0.227***	-0.684***
	(0.051)	(0.010)	(0.026)			
Pupil/Teacher Ratio	34.293	41.249	33.662	-6.956**	0.631	7.587***
	(1.512)	(1.528)	(0.712)			
Pupil/Classroom Ratio	29.642	38.264	30.639	-8.622***	-0.997	7.625***
	(1.450)	(0.945)	(0.651)			
Computers	2.917	4.463	6.365	-1.546***	-3.449***	-1.902***
	(0.264)	(0.233)	(0.517)			
Tap Water	0.990	0.970	0.952	0.019	0.037*	0.018
	(0.010)	(0.009)	(0.012)			
Playground	1.167	1.258	1.191	-0.091*	-0.024	0.067**
	(0.038)	(0.024)	(0.023)			
Library	1.010	1.047	1.010	-0.037*	0.000	0.037***
	(0.010)	(0.012)	(0.006)			
Electricity	1.000	1.036	1.003	-0.036	-0.003	0.032**
	(0.000)	(0.013)	(0.003)			
Toilets per Pupil	0.028	0.022	0.032	0.006**	-0.004	-0.010***
	(0.002)	(0.001)	(0.001)			
Books per Pupil	9.963	3.945	8.668	6.018***	1.294	-4.723***
	(2.002)	(0.294)	(0.555)			
N	96	337	293			

Note: The value displayed for t-tests are the differences in the means across the groups. ***, *, and * indicate significance at the 1, 5, and 10 percent critical level. This table shows descriptive statistics for the schools participating in the voucher experiment as compared to other public schools and non-participating private schools from the same neighborhoods in Delhi. Data comes from the 2011 "District Information System for Education" (DISE) database.

Table 2.4: Descriptive Statistics for Student Outcomes

	mean	sd	count
Test Scores			
Math	0.01	0.98	1354
English	0.05	0.98	1354
Hindi	-0.02	0.99	1354
Non-Cognitive Skills			
Growth Mindest	0.00	1.64	1214
Self-Management	0.00	1.42	1227
Social Awareness	0.01	1.47	1244
Math - Self Rank	38.59	29.83	1353
Hindi - Self Rank	39.07	31.83	1353
English - Self Rank	41.63	31.51	1353
Social Outcomes			
Social Hours w/ Peers	3.92	4.22	1332
Social Hours w/ Peers From Diff. Area	2.14	3.24	1329
Portion Top 5 Friends at Same School	0.78	0.34	1342
Portion Top 5 Friends' Fathers Waged	0.20	0.32	891
Parent Aspirations			
Aspires for college or higher	0.67	0.47	1354
Ideal educational attainment	15.10	2.80	1368
Expected educational attainment	14.13	3.36	1335
School characteristics			
Average Distance to School (KM)	1.10	0.68	1331
Average Class Size	39.14	10.41	1297
Portion Years Sat at Desk	0.95	0.16	1349
Portion Years Very Satisfied With School	0.66	0.42	1349
Education Expenditure	2416.00	3029.73	1387
Teacher characteristics			
Never Hit Respondent	0.47	0.44	1348
Often Hit Respondent	0.15	0.32	1348
Never Hit Others	0.33	0.42	1348
Often Hit Others	0.17	0.34	1348

Note: This table shows descriptive statistics for our main outcomes of interest.

2.4 Empirical Specification

Following our pre-specified model, we estimate the intention-to-treat equation shown in equation 2.1, where y_{it} is the student outcome, T_i is a dummy for the treatment variable, and X_i is a vector of student-level baseline covariates. These baseline covariates include age, gender, baseline monthly income, and a baseline asset index. Four students are missing baseline monthly income data. To avoid dropping these observations, we simply predict income for these four students using the baseline asset index. Calculations for multiple imputations show that one imputation has a 99.72 percent relative efficiency to an infinite number of imputations, thus we only take one draw from this prediction. We also include ward fixed effects to account for the randomization design, and a dummy indicating the endline tracking phase due to the fact that the data collection process spanned longer than expected.

$$y_{it}=\alpha+\beta_1 T_i+\beta_2 X_i+\varepsilon_i \tag{2.1}$$

We also estimate the treatment-on-the-treated effect using a two-stage least squares instrumental variable regression specified in equations 2 and 3, using winning the lottery as an instrument for the portion of the six years of the experiment in which the child attended private school Pr_i.

$$Pr_i=\gamma+\alpha_1 T_i+\alpha_2 X_i+\epsilon_i \tag{2.2}$$

$$y_{it}=\lambda+\beta_3 \widehat{Pr}_i+\beta_4 X_i+\theta_i \tag{2.3}$$

We are unable to distinguish between school quality and peer effects, but note that peer effects have been shown to be small relative to the effects of differences in pedagogy (Duflo et al., 2011). We use randomization statistical inference to test the null hypothesis of no treatment effect, reporting exact p-values in each table. Following Young (2019), we conduct 2000 draws to determine these estimates. To account for multiple hypothesis testing, we use family-wise adjusted standard errors based on 10,000 draws following Westfall, Young, & Wright (1993) along pre-specified groupings. To address concerns regarding attrition, we report upper and lower bounds from Lee (2009) using student gender to tighten the estimates. Table 2.2 presents evidence that attrition did not impact the baseline balance of our sample. Attriters are though slightly more likely to be from the control group and from poorer households. Treatment students were seven percent more likely to be tracked, and students from one standard deviation higher baseline wealth were three percent more likely to be tracked (Table A2.3).

For our primary outcome variables, we estimate an item response theory (IRT) model[17] to generate test scores along three separate academic outcomes: mathematics, English, and Hindi. IRT models are standard in psychometrics and are increasingly common in the development and education literature in economics (see, for instance, (Das & Zajonc, 2010). We estimate hybrid models - two-parameter logistic models for single word reading, three-parameter models for the mathematics questions where a guessing parameter is appropriate, and a partial credit model for the PIRLS items with more than two possible scores. Details on the construction of other outcome variables can be found in Appendix section A2.1.

[17] See Appendix section A2.2 for details of the IRT modeling

2.5 Results

2.5.1 Intent-to-treat estimates - Effect of winning a voucher

Our main results are presented in Table 2.5. Winning the voucher lottery lowers Hindi scores by 0.09 standard deviations, while English and mathematics scores are unaffected. Decomposing the effect on Hindi, we find no effect on reading comprehension, and a reduction in oral reading fluency equivalent to 4 words (from a control group mean of 47 out of 100 words).

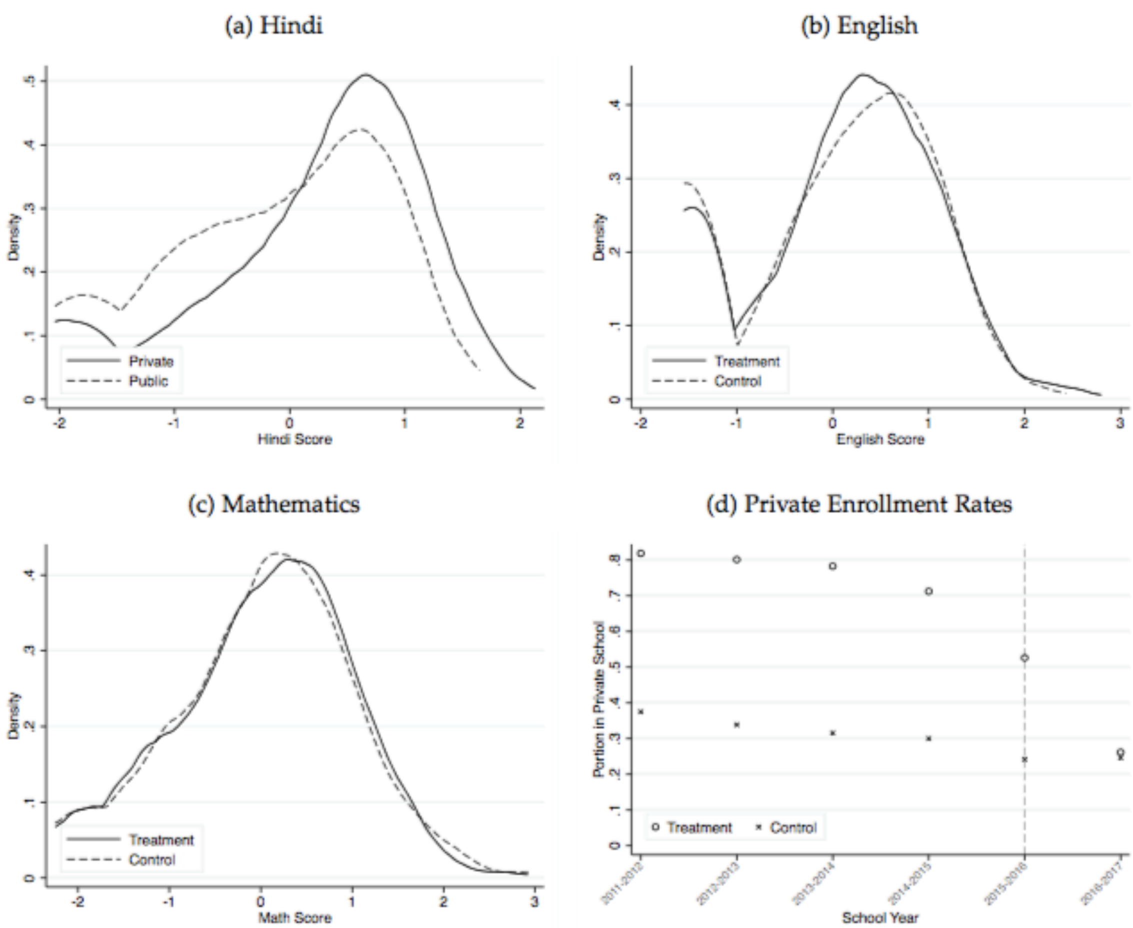
Table 2.5: Voucher Impacts on Academic Learning

	(1) Math	(2) Math	(3) English	(4) English	(5) Hindi	(6) Hindi
Lottery Winner	-0.0258	-0.0388	0.0436	0.0313	-0.0809	-0.0945*
	(0.0567)	(0.0530)	(0.0583)	(0.0523)	(0.0582)	(0.0534)
Baseline controls	No	Yes	No	Yes	No	Yes
Control Mean	-0.000	-0.000	0.000	0.000	0.000	0.000
Exact P-Value	0.646	0.457	0.463	0.540	0.174	0.084
Observations	1367	1367	1367	1367	1367	1367
Clusters	1131	1131	1131	1131	1131	1131
R^2	0.111	0.200	0.078	0.231	0.053	0.190
Lee Bounds	(-0.26, 0.25)	(-0.25, 0.22)	(-0.19, 0.29)	(-0.16, 0.25)	(-0.32, 0.20)	(-0.29, 0.16)

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.
Note: This table shows the impact of winning the voucher lottery on the primary academic learning outcomes. Each outcome is standardised to the control group. The mathematics score is estimated using a three parameter logistic model. The English and Hindi scores are estimated using hybrid item response theory models. Two parameter logistic models are used for the oral reading fluency words, three parameter logistic models are estimated for the dichotomous PIRLS items, and general partial credit models are used for the remaining PIRLS items. Further details on the construction of the outcome variables can be found in Appendix A2.1. Upper and lower bounds from Lee (2009) are reported using student gender to tighten the estimates. Exact p-values show randomization tests based on 2000 draws. Standard errors are clustered at the household level, and all results include an endline phase dummy and ward fixed effects.

Figure 2.1 presents these results visually using kernel density plots. Student Hindi and English performance has a bimodal distribution, highlighting that some students still simply cannot read. One limitation of controlling for the baseline test scores— particularly for Hindi and English—is that because most students could not read at the beginning of the experiment, there is little variation in test scores at the bottom of the distribution. Specifically, 75 percent of students could not read a single Hindi word at baseline, and 84 percent could not read a single English word.

Figure 2.1: Voucher Impacts on Key Education Outcomes



Note: Figures a, b, and c, show the distribution of endline scores for Hindi, English and mathematics, respectively. Solid lines show the treatment group, and dashed lines show the control group. Zero on the x-axis is equivalent to the control group mean. Figure d shows the portion of treatment and control students who attended private school in each given year. The vertical dashed line marks the end of the ENABLE voucher program.

2.5.2 Instrumental variable estimates - Effect of Attending Private School

Next we use the lottery as an instrument for private school attendance. We first note that winning a voucher has a strong effect on private school attendance—more than doubling the amount of time spent at a private school. The voucher increased the portion of the six years between baseline and endline survey spent at a private

Table 2.6: Private School Impacts on Academic Learning: IV Results

	(1) Portion Private School OLS	(2) Math IV	(3) English IV	(4) Hindi IV
Lottery Winner	0.346***			
	(0.0217)			
Portion Private School		-0.121	0.0791	-0.299*
		(0.154)	(0.150)	(0.157)
Control Mean	0.305	-0.000	0.000	0.000
Exact P-Value	0.000	0.963	0.974	0.919
F-Statistic	51.790			
Observations	1387	1354	1354	1354
Clusters	1152	1124	1124	1124
R^2	0.273	0.188	0.242	0.154

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.
Note: This table shows two-stage least squares estimates of the impact of private schooling on test outcomes. We instrument for the endogenous portion of years in private school using the lottery winner status. Column 1 shows the impact of winning the lottery on the portion of years in private school. Each outcome is standardised to the control group. The mathematics score is estimated using a three parameter logistic model. The English and Hindi scores are estimated using hybrid item response theory models. Two parameter logistic models are used for the oral reading fluency words, three parameter logistic models are estimated for the dichotomous PIRLS items, and general partial credit models are used for the remaining PIRLS items. Further details on the construction of the outcome variables can be found in Appendix A2.1. Exact p-values show randomization tests based on 2000 draws. Standard errors are clustered at the household level, and all results include an endline phase dummy and ward fixed effects.

school by 34 percentage points (from a control group mean of 30.5 percent). Though compliance is imperfect, the first stage is strong. The core results on academic outcomes are reported in Table 2.6. The coefficient in Column 4 implies that moving from zero years in a private school to six years reduces Hindi performance by 0.3 standard deviations, with no statistically significant impact on mathematics or English. The effect sizes are similar to those from estimates without the pre-specified controls. The results are robust to looking at raw test scores (not shown) in lieu of the IRT scaled scores.

Our results contrast somewhat with those found by earlier studies, which generally found positive or null results. In particular Wolf et al. (2015) analysis of two-year data from the same experiment found positive effects for all subjects, which were statistically significant only for English (Table 2.7).

Table 2.7: Comparison of estimated private school effects from our study with previous studies

Study	Math	English	Hindi	Telugu	N	Years Exposure	Attrition Rate
Crawfurd, Patel, & Sandefur (2019)	-0.121 (0.16)	0.079 (0.152)	-0.299* (0.158)		1,354	5	13%
Wolf, Egalite, & Dixon (2015)	0.12 (0.173)	0.25*** (0.128)	0.07 (0.151)		1,143	2	29%
Muralidharan & Sundararam (2015)	-0.031 (0.052)	0.229* (0.07)	1.074*** (0.068)	-0.017 (0.051)	4,385 (0.045)	4	18%
Singh (2015) (Urban 8 yos)	0.03 (0.15)				458	3	4%

Note: This table presents a comparison of the estimated private school effects from our study with those from previous studies, including those from Wolf et al. (2015) on the same experiment that we study, after two years, from (Muralidharan and Sundararaman, 2015) on a similar voucher experiment in rural Andhra Pradesh, and from (Singh, 2015) using a value-added model with data from urban Andhra Pradesh.

2.5.3 Robustness

Why might private schools lead to no improvement in outcomes? First, we note that we can rule out that our results are driven by substitution away of parental effort, as has been documented in other contexts (Das, Dercon, Habyarimana, & Krishnan, 2007). We see no change in parent spending or time input due to the voucher (Table 2.8).

Second, winning the voucher lottery could also have direct self-confidence or psychological incentive effects. Such effects would bias our estimates upwards not down, and in any case we have direct measures of self-confidence, finding little effect (Table 2.8).

2.6 Understanding School Choice: Why is Demand So High When Average Impacts Are So Low?

Up to this point, our results present something of a paradox: a majority of households in urban Indian pay to send their children to private school. The offer of private-school vouchers in East Delhi generated excess demand, leading to a lottery system to ration the places. Yet we find no average impact on test scores or most other non-learning outcomes. Why is demand so high for vouchers if returns to private schooling are so low?

Given our main result finding negative or null effects of attending private schools on learning, we investigate three possible explanations for this finding. First, that there may be other non-learning benefits from attending private schools, second that returns vary for different groups depending on what their outside option was, and third that parents face an information asymmetry and incorrectly believe private schools to be better than they are in terms of learning.

2.6.1 Testing for Non-Learning Benefits

First, does private school enrollment yield gains in terms of other outcomes previously held to drive school choice? When parents and students engage in school choice, they are typically maximising over a set of attributes which may include learning and teaching quality, but also reputational concerns, peer quality, and other amenities (MacLeod and Urquiola, 2015).

For the most part, results are equally underwhelming as in the case of academic outcomes. Looking at non-academic student outcomes, we document no impact of winning the lottery on important non-cognitive outcomes including social awareness, growth mindset, or self-confidence (Table 2.8). Social networks also seem to be unaffected by the voucher. The composition of students friendship groups is also unaffected. We find no evidence of voucher impact on teacher absenteeism or on the practice of corporal punishment in the classroom. We do not document any differences in beliefs about child ability from the voucher. Winning the lottery had no impact on parental expectations of that child's future earnings, expected educational achievement, or probability of working in the formal sector. The voucher had no spillovers on other family members academic performance, and in particular, non-participating siblings of voucher winners are no more likely to attend private school (siblings of lottery winners were eligible to also receive vouchers only if they were aged 5-7 years at baseline). If anything, survey responses suggest that receiving a voucher improves parental opinions about the government's capability to provide high quality education in government schools.

We are unable to test with our data the perception that private schooling may improve prospects in the marriage market (Srivastava, 2006), that private schools may improve spoken English if not reading, and that there may be labour market returns to private school attendance and spoken English proficiency (Asadullah, 2009; Azam et al., 2013; Chakraborty and Bakshi, 2016).

Table 2.8: ITT effects on other student, school, and family outcomes

Outcome	β	SE	Control Mean	Exact P-Value	Lee Bounds	N	Clusters
Non-Cognitive							
Growth Mindset	-0.05	0.08	-0.02	.477	(-0.31, 0.32)	1224	1041
Self-Management	-0.14	0.08	0.05	.076	(-0.52, 0.23)	1240	1041
Social Awareness	-0.06	0.08	0.03	.478	(-0.47, 0.23)	1257	1058
Math - Self Rank	3.00	1.45	36.75	.047	(-4.00, 8.57)	1366	1131
Hindi - Self Rank	2.14	1.57	37.50	.176	(-5.12, 8.13)	1366	1131
English - Self Rank	0.95	1.63	40.82	.574	(-6.68, 7.06)	1366	1131
Social Outcomes							
Social Hours w/ Peers	-0.22	0.23	4.03	.333	(-1.37, 0.53)	1344	1117
Social Hours w/ Peers From Diff. Area	-0.21	0.18	2.24	.243	(-1.15, 0.31)	1341	1117
Portion Top 5 Friends at Same School	-0.03	0.02	0.79	.176	(-0.09, 0.07)	1355	1123
Portion Top 5 Friends' Fathers Waged	0.00	0.02	0.20	.855	(-0.13, 0.07)	900	769
School Qualities							
Average Distance to School (KM)	-0.01	0.04	1.10	.896	(-0.18, 0.12)	1337	1112
Average Class Size	-2.98	0.59	40.66	0	(-6.01, -0.41)	1303	1089
Portion Years w/ Desk	0.03	0.01	0.94	.002	(0.01, 0.07)	1355	1125
Portion Years Very Satisfied w/ School	0.03	0.02	0.65	.174	(-0.04, 0.12)	1355	1125
Education Expend.	32.64	162.46	2438	.839	(-835, 528)	1402	1160
Teacher Qualities							
Never Hit Respondent	0.01	0.02	0.46	.729	(-0.08, 0.09)	1354	1124
Often Hit Respondent	-0.01	0.02	0.15	.589	(-0.11, 0.04)	1354	1124
Never Hit Others	0.02	0.02	0.32	.412	(-0.08, 0.09)	1354	1124

Continued on next page

Table 8 – *Continued from previous page*

Outcome	β	SE	Control Mean	Exact P-Value	Lee Bounds	N	Clusters
Often Hit Others	-0.01	0.02	0.18	.466	(-0.12, 0.04)	1354	1124
Post-Voucher							
Attends School	-0.02	0.02	0.93	.296	(-0.06, 0.08)	1402	1160
Attends Private School	0.04	0.02	0.23	.077	(-0.06, 0.12)	1280	1076
Parents Hire Priv. Tutor	0.01	0.03	0.31	.590	(-0.09, 0.09)	1402	1160
Parental Expectations							
Attend College	0.03	0.03	0.65	.318	(-0.05, 0.13)	1366	1130
Ideal Schooling Years	0.06	0.15	15.02	.708	(-0.54, 0.67)	1377	1141
Expected School Yrs	-0.01	0.19	14.02	.969	(-0.73, 0.93)	1343	1117
School 5 Yrs	-0.01	0.02	0.93	.642	(-0.04, 0.09)	1349	1124
School 10 Yrs	-0.03	0.03	0.60	.295	(-0.13, 0.08)	1183	992
Formal Work 5 Yrs	0.00	0.01	0.04	.970	(-0.06, 0.03)	1349	1124
Formal Work 10 Yrs	0.05	0.03	0.26	.099	(-0.07, 0.13)	1183	992
Informal Work 5 Yrs	0.01	0.01	0.03	.457	(-0.04, 0.03)	1349	1124
Informal Work 10 Yrs	-0.02	0.02	0.13	.456	(-0.13, 0.04)	1183	992
Income 5 Yrs	-148.94	105.37	408	.129	(-555, 88)	1171	984
Income 10 Yrs	-388.41	592.97	5412	.518	(-3422, 1219)	1228	1030
Sibling Spillovers							
Portion In School	-0.03	0.02	0.88	.123	(-0.08, 0.08)	1245	1018
Portion In Priv. School	-0.02	0.02	0.23	.429	(-0.14, 0.05)	1173	959
Sibling's Hindi	-0.04	0.03	0.50	.263	(-0.17, 0.09)	943	873
Sibling's Math	-0.04	0.03	0.33	.175	(-0.19, 0.05)	943	873
Sibling's English	-0.07	0.05	0.02	.207	(-0.33, 0.10)	1402	1160

Continued on next page

Table 8 – *Continued from previous page*

Outcome	β	SE	Control Mean	Exact P-Value	Lee Bounds	N	Clusters
Parental Help on Homework	0.05	0.29	4.57	.867	(-1.42, 1.04)	1164	953
Private Tutor Expend.	15.84	20.55	180	.501	(-92, 76)	1102	914
Education Expenditure	-63.60	157.08	2344	.698	(-877, 430)	1262	1073
Parental Beliefs							
Govt. Economy Duty	-0.07	0.06	-0.00	.211	(-0.24, 0.21)	1399	1159
Govt. Education Duty	-0.10	0.06	-0.00	.088	(-0.27, 0.20)	1399	1159
Govt. Capability	0.10	0.06	-0.01	.096	(-0.21, 0.28)	1399	1159
Parent Outcomes							
Can Read Hindi	-0.02	0.03	0.47	.465	(-0.12, 0.07)	1399	1159
Can Read English	-0.03	0.02	0.19	.218	(-0.13, 0.03)	1399	1159
Cantril Ladder	-0.16	0.12	5.16	.172	(-0.75, 0.35)	1385	1147

Note: Each row shows the ITT effect of winning a voucher on a separate outcome variable. Upper and lower bounds from Lee (2009) are reported using student gender to tighten the estimates. Exact p-values show randomization tests based on 2000 draws. Standard errors are clustered at the household level, and all results include an endline survey-round dummy and ward fixed effects.

2.6.2 Heterogeneous Effects: Impacts May Vary by Outside Options

In this section, we explore the possibility that tension between high demand and low average impact can be explained by heterogeneity in household's outside options. We begin by laying out a simple conceptual framework that demonstrates how school choice in the face of a voucher for low-cost private schools may yield null impacts on learning even if the quality of these schools is superior to government schools. This argument motivates estimation of a specific form of heterogeneous effects in terms of both take-up and voucher impacts.

Conceptual framework

Suppose households select a school to maximize utility as a function of school quality and other consumption, $U(Q_j, C)$ subject to a budget constraint defined by their wealth, *W*. They choose from three types of schools, j $\in$ {g,l,e}: free government schools, *g*; low-cost private schools that (by hypothesis) provide at least minimally higher quality, *l*; and elite schools that provide higher quality at yet higher cost, *e*. In short, qualities and prices are ranked such that $Q_g < Q_l < Q_e$ and $0 = P_g < P_l < P_e$. With standard assumptions on the shape of the utility function, the poorest households will opt for government schools, the richest into elite private schools, and some intermediate households will send their children to low-cost private schools.

In this simple setup, the introduction of a voucher that sets $P_l = 0$ will induce poor households to switch from government to low-cost private schools, with some improvement in education quality. In addition, however, vouchers will induce some wealthy households to switch from elite schools into lower-quality low-cost private schools.

The implication for the empirical analysis is that if quality and price of school types are ranked in the way we posit here, we should find that (i) the take-up of vouchers is declining in wealth, (ii) the poorest households who spend the least on private education in the absence of a voucher obtain the largest academic benefits from their introduction, and (iii) some wealthier households who are likely to send their children to elite schools otherwise, may suffer negative impacts on academic outcomes from the introduction of the voucher.

Heterogeneous Effects by expected spending on school

First, using data from the control group, we estimate the amount spent on schooling $\hat{y}_i$ as a function of baseline covariates, $\mathbf{X}_i$, including household income, wealth, religion, caste, and the child's age, gender, and baseline test scores. We then use these covariates to predict expected spending on school for the treatment group.

$$\hat{y}_i = \mu + \phi X_i + u_i \tag{2.4}$$

We predict $\hat{y}_i$ using Ordinary Least Squares, and five different machine learning algorithms; LASSO, Random Forest, BOOST, Support Vector Machines, and Kernel-Based Regularized Least Squares (KRLS). By far the best performing predictor is the Random Forest, with an out of sample R-squared of .286, compared to an R-squared of .128 for Boost (second best) and .051 for OLS (see Table A2.4).[18] Selecting the random forest estimator as the best predictor of actual spending, we then categorise children into three groups based on their predicted spending. For students in the control group who never attended a private school, median spending was 2,500 rupees. We therefore assign anyone in either treatment or control group with predicted spending below 2,500 to be in the free government school group. Anyone with predicted spending above the 7,300 rupees that is the maximum value of the ENABLE voucher is assigned to the elite private school group. The remaining students are assigned as those who might have attended a similar low-cost private school anyway. There are 1,054 children in the predicted free government school group (65 percent), 136 children

[18] This ordering is consistent with (Mullainathan & Spiess, 2017)

in the predicted low-cost private school group (8 percent), and 428 children in the predicted higher-cost private school group (26 percent).

Next, we test for heterogeneous treatment effects by interacting the indicator for winning the lottery with each child's predicted school type from equation (4).

$$y_{it}=\alpha+\beta_1 T_i+\beta_2 T_i \times \Pr(\widehat{Gov})_i+\beta_3 T_i \times \Pr(\widehat{Elite})_i+\beta_4 X_{it-1}+\varepsilon_i \qquad (2.5)$$

The results of this analysis is consistent with the conceptual framework - we see some negative treatment effects for those predicted to attend a higher fee private school, and some positive effects for those predicted to attend free government schools.

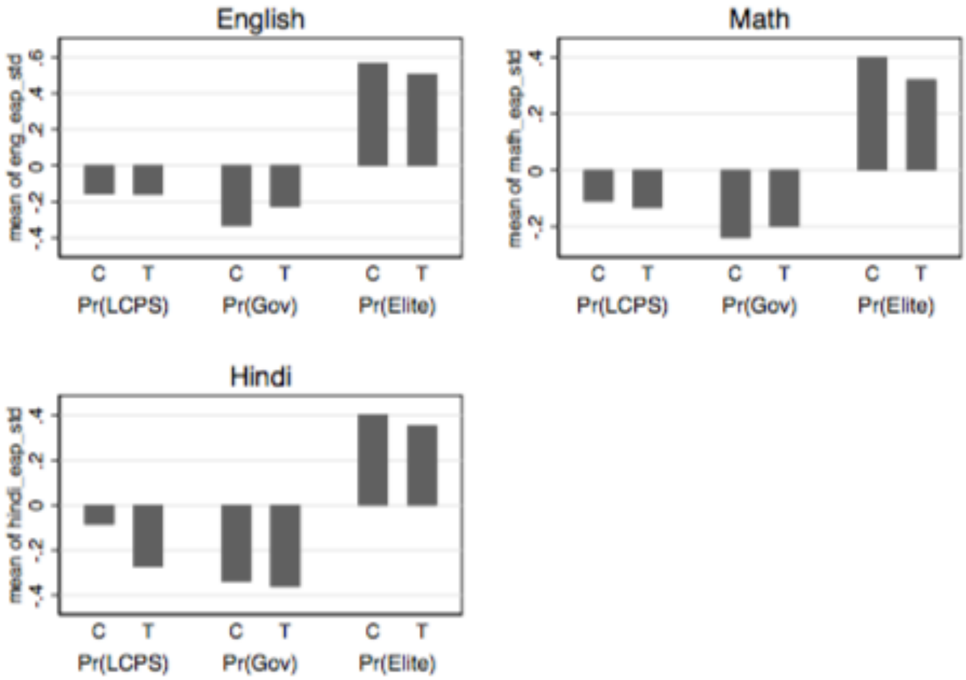
Table 2.9: Heterogeneous Voucher Impacts

	(1) Used voucher	(2) Class size	(3) English	(4) Math	(5) Hindi
Lottery Winner	0.542***	-3.279***	0.0253	-0.0441	-0.158**
	(0.0865)	(0.763)	(0.0681)	(0.0720)	(0.0732)
Lottery winner	0.147	-0.0665	0.270	0.221	0.388*
× Predicted Government	(0.144)	(3.314)	(0.186)	(0.212)	(0.222)
Lottery winner	-0.148	1.435	-0.221**	-0.137	-0.0181
× Predicted Elite Private	(0.160)	(1.279)	(0.111)	(0.111)	(0.112)
Control Mean	0.211	40.665	-0.000	-0.000	-0.000
Exact P-Value	0.000	0.000	0.610	0.390	0.004
Observations	748	1303	1367	1367	1367
Clusters	611	1089	1131	1131	1131
R^2	0.197	0.109	0.270	0.217	0.218

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.

Note: This table shows heterogeneous ITT effects of winning the voucher lottery on voucher take-up, school quality measures, and the primary academic learning outcomes. Each outcome is standardised to the control group. The mathematics score is estimated using a three parameter logistic model. The English and Hindi scores are estimated using hybrid item response theory models. Two parameter logistic models are used for the oral reading fluency words, three parameter logistic models are estimated for the dichotomous PIRLS items, and general partial credit models are used for the remaining PIRLS items. Further details on the construction of the outcome variables can be found in Appendix A2.1. Exact p-values show randomization tests based on 2000 draws. Standard errors are clustered at the household level, and all results include an endline phase dummy and ward fixed effects.

Figure 2.2: Mean outcomes by Predicted School Type



Note: Figure shows average endline test scores for the three groups of predicted school type.

2.6.3 Asymmetric information

Our third hypothesis is that parents can't observe true school quality (value-added), and at best observe average test scores or peer quality and indicators of school infrastructure (Abdulkadiroğlu, Pathak, & Walters, 2018). This theory is consistent with evidence from various contexts showing that providing parents with information about comparative school performance improves choices (Hastings and Weinstein, 2008; Andrabi et al., 2017; Afridi et al., 2018). First, we compare estimates of school test scores and value-added with parental preferences.

We estimate the quality of schools that voucher winners selected based on their value-added. We incorporate school fixed effects into equation 2.1. We do this in two ways. First, we interact treatment with the school allocated in the lottery. Second, we additionally scale these coefficients by the number of years the student actually attended the school.

Given the small sample sizes for some of these schools, we risk estimating particularly extreme value-added when the estimation error is high. To address this bias-variance tradeoff, we implement an Empirical Bayes shrinkage estimator from Morris (1983) that has become common in the education and health literature. Following the procedure of Chandra et al. (2016), we shrink the school value-added estimates toward the mean of the true distribution. Further details are discussed in Appendix A2.3. Since most students were allocated to their first-choice school, we do not have enough observations to use the "overflow" students to causally estimate the marginal returns to a given school. Most school-choice in this context is determined by location, so our specification relies on the assumption that conditional on the neighborhood fixed effects, household income and wealth, gender, and age controls that we use, the allocation of one private school versus another is based on idiosyncratic variation in the geographic distribution of students and households that is orthogonal to value-added.

We use data on parental preferences from their ranking of schools. All lottery winners were asked to rank their top five schools from the full list of 105 schools. Whilst estimates are imprecise due to the small number of schools (78), we do show that the relationship between parent rankings and average test scores is much stronger ($R^2 = 0.087$) than the relationship between preferences and school value-added ($R^2 = 0.006$). Other school characteristics are not statistically significantly correlated with school rankings.

Table 2.10: School Quality and Parental Preferences

	(1)	(2)
School Value-Added (English)	-0.0365	
	(0.405)	
School Value-Added (Math)	-0.0578	
	(0.181)	
School Value-Added (Hindi)	0.181	
	(0.439)	
Average test scores (English)		-0.0115
		(0.382)
Average test scores (Math)		-0.334**
		(0.150)
Average test scores (Hindi)		0.571
		(0.459)
Observations (Students)	588	588
Clusters (Schools)	78	78
R^2	0.007	0.095

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.
Note: This table shows the correlation between parental preferences and school characteristics. The dependent variable in each case is the mean rank for each school given by parents of treatment students following the initial lottery. Each parent was asked to rank their top 5 school choices from the full list of 105 schools. 1 is the best rank and 5 is the worst. Standard errors are clustered at the school level.

Second, if we expect parents to learn more about true (lower than expected) quality over time, we would expect voucher take-up to decline over time, and to decline the most for those in the weakest voucher schools. Figure A2.1 shows these trends separated by wealth tertile—which are dramatic. Take-up fell steeply from more than three quarters in the first year to 41 percent in year five. As Figure 2.1d has shown already, take-up falls at a similar rate for treatment and control students. Richer households had lower voucher take-up overall, though parents across the wealth distribution tended to abandon the voucher program at a similar rate. The rapid convergence in year six is most likely due to budget constrained households being unable to afford private schools, though we cannot distinguish that effect from adjustments stemming from learning about school quality.

To further understand the reasons for moving from private to public schools, we estimate which baseline characteristics correlate with remaining in private school at endline. Overall one third of students who were in a private school at the start of the programme had moved to a public school by the end (including both treatment and control students). By contrast, just five percent of students who started in a public school ended in a private school (Table A2.5).

Unsurprisingly, those we predict (based on baseline student characteristics) to spend the most on school, are more likely to be in private school at endline, whether they started in government or private school. For those who started in private schools, choosing a school with larger class sizes or without desks is associated with being less likely to remain in private school. For those who started in a government school, satisfaction with the initial school is associated with being less likely to be in private school at endline. Average school test scores and distance from home to school are not correlated with endline private school enrolment for either group (Table 2.11). This is somewhat consistent with the theory that parents do not perfectly observe school quality ex ante, but do gain information over time, though neither raw test scores nor value-added are associated with endline take-up.

Thirdly, Table 2.8 shows that on key visible inputs—specifically, classroom desks and class size—voucher recipients did see modest but statistically significant gains over control.

This raises the possibility that parents overestimate the importance of these inputs, or confuse correlation and causation when assessing the learning outcomes implied by private education. They could be forgiven, as our data suggest common methods to estimate private school effects using non-experimental data significantly overstate the returns in this context. Specifically, we compare estimated private school effects using an OLS-model with contemporaneous covariates for family background, a value-added model using baseline student test scores, and lower bounds on these estimates accounting for selection on unobservables proportional to the estimated selection on observable covariates as suggested by (Oster, 2016). All three methods produce positive estimates (Figure A2.4), and for the case of English and Hindi – but not mathematics – the non-experimental estimates fall outside the 95% confidence interval of our experimental estimates.

Table 2.11: Correlates of Remaining in Private School at Endline

	(1) Began in Gov	(2) Began in Priv
Lottery Winner	0.0742**	0.0309
	(0.0339)	(0.0446)
Predicted Government	-0.0470*	-0.147
	(0.0265)	(0.0995)
Predicted Elite Private	0.0459	0.0912**
	(0.0361)	(0.0439)
Average test scores (Standardised)	-0.0231	0.0140
	(0.0274)	(0.0182)
Very Satisfied with School (1st School)	-0.0576**	-0.0227
	(0.0264)	(0.0460)
Distance to school (Standardised)	0.00637	0.00949
	(0.0124)	(0.0210)
Class-size (Standardised)	-0.0178	-0.0782***
	(0.0116)	(0.0227)
Did not have a desk	-0.0271	-0.416***
	(0.0286)	(0.106)
Observations	425	606
Clusters	383	529
R^2	0.114	0.121

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.

Note: This table shows which baseline student and school characteristics correlate with remaining in private at endline. Standard errors are clustered at the household level.

2.7 Conclusion

We estimate the impacts of a voucher lottery for low-cost private schools among poor households in Delhi, shedding light on the likely effects of a sea change in education in urban India. We show that receiving a voucher for five years of private schooling lowered Hindi scores, with no impacts on mathematics or English. Examining alternative drivers of the high demand for private schooling, we provide evidence against several common hypotheses: parents do not use vouchers to enroll students in schools closer to home, they do not choose schools that generate greater non-cognitive skills or lower corporal punishment, they experience only slightly improved classroom facilities, and no discernible change in social networks. Among voucher winners, private school enrollment declined significantly over time, and particularly for those who chose schools with large class sizes. This is consistent with the notion that school quality is unobservable ex ante, but parents learn about quality over time. We find some evidence that the null effects can be explained by substantial heterogeneity. Those who would have otherwise attended a free government school do see some positive impacts, and those who would otherwise have attended a higher-cost private school see larger negative impacts.

A key question in evaluating our results across various dimensions is whether they constitute an informative null result, or leave open the possibility of economically significant returns. Although we improve substantially on the 29 percent attrition observed at the two-year follow-up survey (Wolf et al., 2015), our attrition rate of 13 percent is nonetheless less than ideal and limits our ability to rule-out some positive effects. As in any experiment of this kind, we can only speak to the impacts on compliers, and cannot rule out the possibility that students already enrolled in private schools (‘always-takers’) experience higher returns. Our sample and therefore focus includes only low-cost private schools in one neighborhood of Delhi. Future research could usefully test our findings in other settings, as suggested by at least some observational data analysis (Moore, 2017). Our focus is also only on voucher students and not on other remaining students in government schools (studied by Muralidharan and Sundararaman, 2015) or on prior students in private schools (studied by Rao, 2018). We are not able to comment on whether voucher students may face discrimination, making their results unrepresentative of other students in low-cost private schools.

What implications do our results have for the continued government scale-up the Right to Education Act Section 12(1)(c)? Though the policy was designed to give low-income students access to better quality private schools, our results suggest that a large proportion of lower-fee private schools fail to provide such quality. From a policy perspective, however, it bears noting that schools in our sample generate roughly equivalent learning at a per pupil expenditure that is roughly half of nearby government schools. Whether the expansion of government-funded places in low-cost private schools leads government to reduce spending on government schools, or what the equilibrium effects of such a policy might be, remain open questions.

There are also important differences between the voucher experiment evaluated in this study and the RTE policy. First, the RTE policy reserves 25 percent of seats in private schools for eligible low-income students. In none of the schools in our experiment did the voucher children take up this many seats. Second, the schools in our experiment were able to expand their number of places without turning away any other fee-paying students, which might not be the case for all schools. Allende, Gallego, & Neilson (2019) for example highlight the importance of supply response in determining the aggregate welfare implications of school choice. Third, the relatively small number of vouchers funded through our experiment means that there are unlikely to have been substantial spillover effects on students left behind in government schools (as for example considered by Muralidharan & Sundararaman, 2015). This could be an important mechanism in the policy at scale. Overall, the precise design of the RTE policy is left to individual states and therefore has substantial variation. Our study speaks most directly to inferences about the overall existing quality of low-cost urban private schools, rather than to possible broader or general equilibrium effects.

Government also bears some duty to students who attend private schools through entirely private finance. Given our results consistent with imperfect information, a plausible role for government could be providing the public good of better information about school quality, as has shown to be effective at improving private school quality in the US (Hastings and Weinstein, 2008), India (Afridi et al., 2018) and Pakistan (Andrabi et al., 2017), or focusing on private school market failures in other areas (Andrabi, Das, & Khwaja, 2015). Since our endline survey, the South Delhi government has announced that government schools will shift to English medium–a move provoked by the popularity of English-medium private schools.[19] The evidence presented here suggests that this change may have little effect on the English reading ability of students in government schools.

[19] https://timesofindia.indiatimes.com/city/delhi/south-corp-schools-to-teach-in-english/articleshow/59905602.cms

Appendix A2.1: Variable Construction

The construction of our outcome variables follows the methods pre-specified before the endline results were collected. Table notes include descriptions of many variables. Total education expenditure is the sum of the amount spent on tuition fees, uniforms, textbooks and school materials, transportation to school, other school fees, and private tuition.

The non-cognitive outcomes in Table 2.8 were calculated based on the first component of principal component analyses (PCA) of a series of Likert scale questions. Students chose among completely true, mostly true, somewhat true, a little true, and not true at all. The growth mindset variable is based on: “My intelligence is something that I can’t change very much”; “Challenging myself won’t make me any smarter”; “There are some thing I am not capable of learning”; and “If I am not naturally smart in a subject, I will never do well in it.” Similarly, the self-management variable is based on: “I get distracted easily”; “I refuse things that are bad for me, even if they are fun”; “I do thinks that feel good in the moment but regret later on”; and “Sometimes I can’t stop myself from doing something, even if I know it is wrong.” Student responses to “I refuse things that are bad for me...” were reversed before the PCA was conducted. Social awareness was based on the following: “I get along with students who are different from me”; “When others disagree with me, I am respectful of their views”; “I can disagree with others without starting an argument”; “I care a lot about other students feelings.”

The political economy beliefs in Table 2.8 were constructed by the same method. The options for parents included strongly agree, somewhat agree, somewhat disagree, and disagree. The government capability index is constructed from “The government is capable of delivering high quality education” and “The public school that is closest to your home is a good school.” The government education duty index is based on “The government has a duty to provide free high quality education to everyone” and the inverse of “Parents should take more responsibility to ensure that their children get a high quality education.” The government economy duty index is formed from “When it comes to the economy, the government has a duty to ensure that everyone has a job” and the inverse of “When it comes to the economy, people should take responsibility to ensure that they find themselves a job.”

We constructed school-characteristic variables from the DISE data.

Appendix A2.2: Item Response Theory Models

We estimate test scores using techniques from Item Response Theory. Let i index students and j index items (test questions.) The English and Hindi oral reading fluency (ORF) assessments—word lists which do not have multiple choice options—lend themselves to a two-parameter logistic model (2PL) as shown in equation 2.6. The mathematics assessment, by contrast, includes multiple choice options, so we expand the model to a three-parameter one (3PL) by adding a guessing parameter c_j as shown in equation 2.7. For the polytomous items from the Progress in International Reading Literacy Study (PIRLS), we scale scores with a generalized partial credit model (GPCM) using equation 2.8, where l_i is student i's grade on a scale with m_j categories (in our case, $m_j = 3$.) To construct the primary outcome variables in our specification, we use a hybrid model of 2PL, 3PL, and GPCM, producing a single B_i for each student. We estimate a 2PL model for the English ORF items administered to siblings.

$$\Pr(X_{ij} = 1 | \theta_i, a_j, b_j) = \frac{\exp[a_j(\theta_i - b_j)]}{1 + \exp[a_j(\theta_i - b_j)]} \tag{2.6}$$

$$\Pr(x_{ij} = 1 | \theta_i, a_j, b_j, c_j) = c_j + (1 - c_j)\frac{\exp[a_j(\theta_i - b_j)]}{1 + \exp[a_j(\theta_i - b_j)]} \tag{2.7}$$

$$\Pr(x_{ij} = l | \theta_i, a_j, b_j, d_{j,1}, \ldots, d_{j,m_j-1}) = \frac{\exp\left(\sum_{v=0}^{l_i} a_j\left(\theta_i - b_j + d_{j,v}\right)\right)}{\sum_{g=0}^{m_j-1} \exp\left(\sum_{v=0}^{g} a_j\left(\theta_i - b_j + d_{j,v}\right)\right)} \tag{2.8}$$

Appendix A2.3: Heterogeneous effects by gender

On average, girls score higher in both English and Hindi than boys, but unlike Wolf et al. (2015) who found larger positive effects for girls, we do not observe any differential treatment impacts by gender on our main learning outcomes, or on the likelihood of attending a private school (Table A2.1). We do though see differential impacts by gender for our main non-cognitive outcomes. Male lottery winners see substantially worse results (than male control students) for growth mindset, self-management, social awareness, and self-confidence (self-perceived skill rank in Mathematics), but higher life satisfaction (by 0.4 points on a 1 - 10 Cantril ladder). Female lottery winners see none of these changes (Table A2.2).

Table A2.1: Heterogeneous Voucher Impacts on Academic Learning by Gender

	(1) Math	(2) English	(3) Hindi
Lottery Winner	-0.0264	0.0414	-0.136*
	(0.0722)	(0.0716)	(0.0755)
Treatment × Female	-0.0261	-0.0212	0.0879
	(0.0984)	(0.0958)	(0.0989)
Control Mean	-0.000	0.000	0.000
Exact P-Value	0.611	0.427	0.013
Observations	1367	1367	1367
Clusters	1131	1131	1131
R^2	0.200	0.231	0.190
Lee Bounds	(-0.24, 0.18)	(-0.27, 0.18)	(-0.24, 0.30)

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.
Note: This table shows the impact of winning the voucher lottery on the primary academic learning outcomes. Each outcome is standardised to the control group. The mathematics score is estimated using a three parameter logistic model. The English and Hindi scores are estimated using hybrid item response theory models. Two parameter logistic models are used for the oral reading fluency words, three parameter logistic models are estimated for the dichotomous PIRLS items, and general partial credit models are used for the remaining PIRLS items. Further details on the construction of the outcome variables can be found in Appendix A2.2. Upper and lower bounds from Lee (2009) are reported using student gender to tighten the estimates. Exact p-values show randomization tests based on 2000 draws. Standard errors are clustered at the household level, and all results include an endline phase dummy and ward fixed effects.

Table A2.2: Heterogeneous Voucher Impacts on Non-Cognitive Student Outcomes

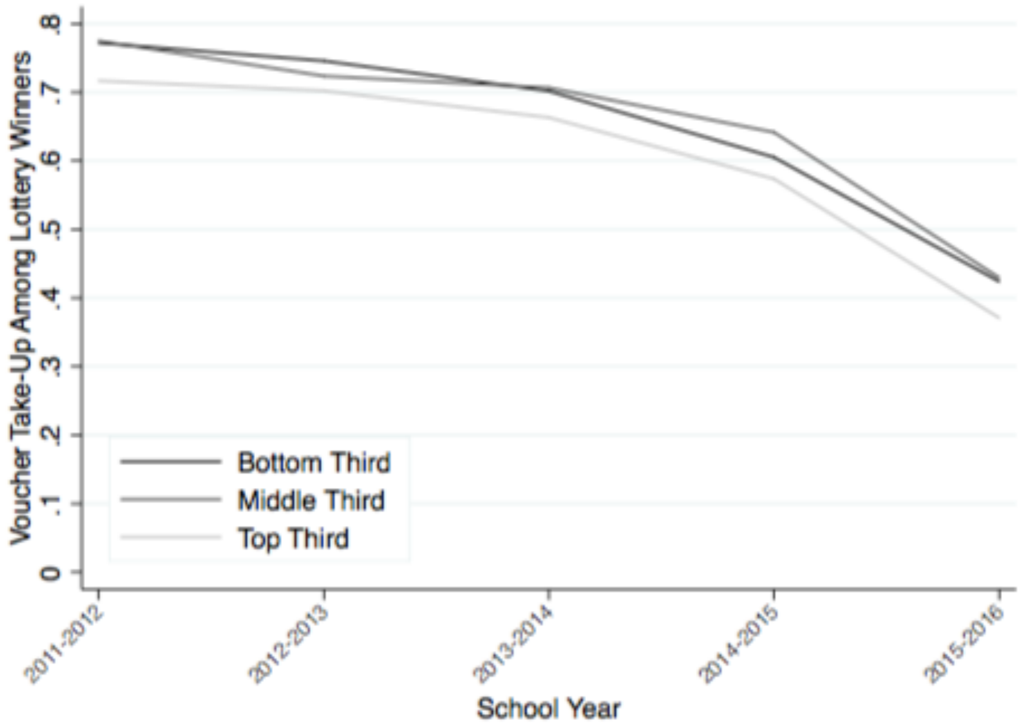
	(1) Growth Mindset	(2) Social Awareness	(3) Self Mgmt.	(4) Math Self Rank	(5) Hindi Self Rank	(6) English Self Rank
Lottery Winner	-0.213**	-0.206*	-0.185*	4.353**	3.230	1.647
	(0.106)	(0.109)	(0.109)	(1.986)	(2.147)	(2.188)
Treatment × Female	0.329**	0.132	0.268*	-2.850	-2.291	-1.468
	(0.149)	(0.157)	(0.155)	(2.887)	(3.121)	(3.069)
Age (months)	-0.00493	-0.00223	0.000860	0.0493	0.0537	0.0205
	(0.00400)	(0.00418)	(0.00392)	(0.0773)	(0.0859)	(0.0849)
Female	-0.0965	0.114	-0.251**	-0.197	-0.253	-2.589
	(0.108)	(0.115)	(0.114)	(2.098)	(2.296)	(2.273)
Baseline Monthly Income (1000s)	-0.0202	0.0244	-0.0141	-0.305	-0.448	-0.393
	(0.0149)	(0.0224)	(0.0151)	(0.290)	(0.286)	(0.305)
Baseline Asset Index	-0.00918	-0.0224	0.0165	0.201	-0.0888	0.0147
	(0.0297)	(0.0320)	(0.0298)	(0.647)	(0.669)	(0.650)
Control Mean						
Exact P-Value						
F-Statistic	1.9	2.0	2.5	3.0	3.5	1.9
Observations	1224	1240	1257	1366	1366	1366
Clusters	1041	1041	1058	1131	1131	1131
R^2	0.418	0.121	0.178	0.238	0.237	0.220

by Gender

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses. Note: This table shows the impact of winning the voucher lottery on non-cognitive student outcomes. The growth mindset, social awareness, and self-management outcomes are principal component analysis results of a series of questions, detailed further in Appendix A2.1. The mathematics, Hindi, and English self ranks are students' responses to, "Consider other students like you who are the same age and who live in the same place as you. If you were ranked against 100 students from this group based on your [mathematics/Hindi/English] ability, where would you rank? (where 1 is the best and 100 is the worst)" Upper and lower bounds from Lee (2009) are reported using student gender to tighten the estimates. Exact p-values show randomization tests based on 2000 draws. Standard errors are clustered at the household level, and all results include an endline phase dummy and ward fixed effects.

Appendix A2.4: Additional Figures

Figure A2.1: Voucher Take-Up by Wealth Tertile



Note: Figure A2 1 shows average voucher take up over time among lottery winners separated into three wealth tertiles constructed from the first principal component of baseline assets

Figure A2.2: ENABLE Vouchers

Note: This figure shows the vouchers distributed to lottery winners.

Figure A2.3: Sample Math Question

783 can be written as 7 Hundreds + 83 Ones.

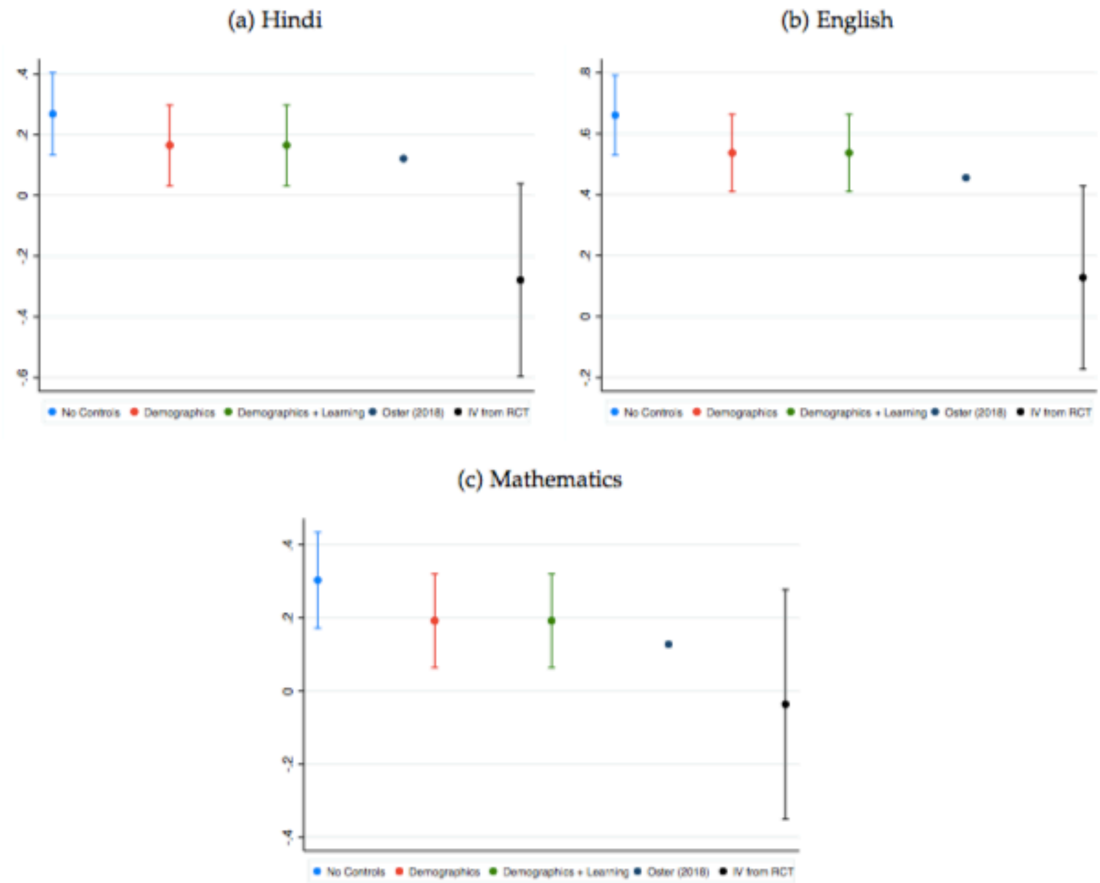
Then, 2041 = ... Hundreds + 41 ones

783 को 7 सैकड़ा + 83 इकाई की तरह लिखा जा सकता है.

तो, 2041= ...... सैकड़ा + 41 इकाई

Note: This figure shows a sample mathematics question administered to the students at endline.

Figure A2.4: Diminishing Private School Effects



Note: These figures show estimated private school effects in 1) a model with no controls, 2) with baseline controls for age, sex, and family wealth, 3) with baseline controls for age, sex, family wealth, and test scores, 4) With selection on unobservables proportional to estimated selection on observables, 5) instrumental variable estimates using voucher assignment for experimental variation in private school attendance.

Appendix A2.5: Additional Tables

Table A2.3: Predictors of Attrition

	(1) Tracked
Lottery Winner	0.0683***
	(0.0190)
Age (months)	-0.000320
	(0.000762)
Female	0.0139
	(0.0170)
Baseline Monthly Income (1000s)	-0.00235
	(0.00405)
Standardized Asset Index	0.0317***
	(0.0100)
F-Statistic	4.772
Observations	1,618
Clusters	1,338
R^2	0.057

* p <0.1, ** p <0.05, ** * p <0.01. Standard errors are shown in parentheses.
Note: This table shows which baseline characteristics correlate with being tracked in the endline survey. Standard errors are clustered at the household level, and the model includes ward fixed effects

Table A2.4: Performance of algorithms in predicting school spending

	Training sample (control)	Hold-out sample (treatment)
Random Forest	.963	.286
KRLS	.108	.076
OLS	.133	.051
LASSO	.071	.023
Boost	.371	.128
SVM	.03	.014

Note: This table presents the R-squared from bivariate regressions of the actual value of average household spending on school, on the predicted value. Covariates used to predict average spending are student age, sex, baseline household income and wealth, baseline test scores, religion, caste, parental education, and survey tracking phase.

Table A2.5: Movement between school types over 5 years

	Treatment		*Control*	
	Number	Percent	Number	Percent
Private to Private	384	51.82	152	23.53
Private to Gov	188	25.37	80	12.38
Gov to Gov	148	19.97	401	62.07
Gov to Private	21	2.83	13	2.01
Total	741	100.00	646	100.00

Note: This table categorises all students according to the school type they started in Year 1 of the program and ended in Year 5 of the program.

Chapter 3: Contracting Out Schools at Scale: Evidence from Pakistan

3.1 Introduction

Private school enrolment in low- and middle-income countries more than doubled between 2000 and 2015. This represents growth of five percent per year (compared to 0.2 percent per year in government schools).[20] Though private schools may offer some benefits over government schools, they are typically not accessible to the poorest due to higher fees. Some governments have decided to use public finance to allow poor students to get the (perceived) benefits of private schools.

Can public-private partnerships (PPPs) provide equitable access to cost-effective privately managed schools? And if so what type of arrangements are most effective? One distinction is between public schools that are privately managed (such as US Charter schools or UK Academies), and private schools that receive public funds through a subsidy or voucher programme. Evidence from high-income countries suggests that privately-managed public schools are more promising than vouchers for private schools (Epple et al., 2017; Epple et al., 2016). In developing countries the evidence is more positive on vouchers and subsidies (Aslam et al., 2017; Shakeel et al., 2016), but there are far fewer studies. As yet there is no rigorous study on a large-scale charter-style public school

[20] World Bank World Development Indicators

management programme in a low or middle-income country. This matters as we might not expect findings from high income contexts to hold in lower-income countries in which the capacity of the state to provide effective procurement and regulation are likely to be particularly low. One study looks at the Partnership Schools for Liberia (PSL) pilot that involved 93 schools in its first year. A randomised evaluation found positive effects on learning, but with high and potentially unsustainable costs, and from a sample of schools that were relatively easier to access than the average school in the country (Romero, Sandefur, & Sandholtz, 2017). Two studies have looked at the Colombia Colegios en Concesión (CEC) programme that involved 25 schools in Bogotá. Both suggest that these schools do outperform traditional public schools in test scores, driven by a longer school day (Bonilla-Angel, 2011; Termes et al., 2015). Another study looks at the 'Adopt a School' programme in Pakistan, a precursor of the programme I study in this paper. This programme involved around 1,000 schools in Punjab and 500 schools in Sindh. The study though presents only descriptive analysis without making claims for causal inference (Malik et al 2015). In high-income countries, there is a larger literature on similar programmes, which have grown to a total of 7,000 charter schools in the US over 26 years (D. Epple et al., 2016), and around 5,000 academies in the UK over 16 years (Eyles, Machin, & McNally, 2017).

In this study we provide the first estimates of a large-scale "contract management" public-private partnership in a developing country. The Punjab Public School Support Programme (PSSP) involved the largest ever contracting out of government schools to private management in a single year. Management of 4,276 "failing" government schools was contracted out to a range of non-profit educational organisations and experienced individuals. The introduction of the programme coincided with a four percent year on year increase in total school enrolment in the province, or more than 450,000 students. This includes an increase of 57,514 students in the first converted schools, which represents an increase of 78 percent. We estimate the effect of school conversion on enrolment and test scores with a difference-in-difference estimator, comparing early converters to later converters. We find a large increase in enrolment relative to comparison schools. This increase is concentrated in Katchi grade (Kindergarten), suggestive of new enrolment of previously out of school children. Due to the large change in the composition of students in schools, and the short time frame within which test scores are measured, we do not necessarily expect to see any change in test scores. Test scores decline after conversion, though we are also less confident in this result as prior trends are not parallel. We present some evidence that the drop in exam scores may be due to the entry of new lower performing students. For schools that did not increase their number of exam candidates, there was no change in test scores.

3.2 Schools in Punjab

Punjab is the largest of the four provinces in Pakistan, with a population of 110 million people. Education service delivery is largely decentralised from the federal government to the provincial governments. Like many low- and middle-income countries, Pakistan faces a learning crisis. Nearly half (43 percent) of Grade 3 children in rural Punjab are unable to read a sentence in Urdu aimed at Grade 2 students (ASER Pakistan, 2017). Only 73 percent of primary-aged children (6-10 years old) are in primary school, with a further 11 percent in pre-primary and 16 percent never having attended any kind of school (Table 3.1).

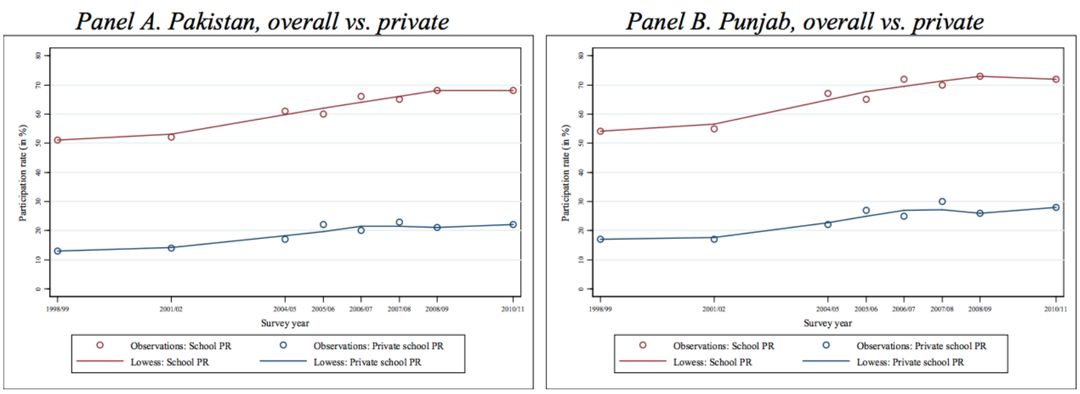
Table 3.1: Enrolment by Class, Thousands, Children Aged 6-10, Punjab

	Government School	Non-State School	Total	Percentage
Never attended			1,926	0.16
Less than 1	626	671	1,297	0.11
Class 1	1,541	1,258	2,799	0.23
Class 2	1,311	1,014	2,324	0.19
Class 3	941	804	1,745	0.14
Class 4	602	499	1,100	0.09
Class 5	313	326	639	0.05
Class 1 - 5	4,707	3,901	8,608	0.71
Class 6 +	130	128	258	0.02
Total			12,088	1.00

Note: Data from the Pakistan Social and Living Standards Measurement Survey 2014-15

Private schooling began to expand in Pakistan in the 1990s – the share of private schooling in total enrolment doubled from 15 to 30 percent between 1991 and 2001 (Andrabi, Das, & Khwaja, 2008), and continued to rise through 2011 (Figure 3.1). Teachers at private schools are more likely to be female, paid less than government teachers, and have less security of tenure (Andrabi et al., 2008).

Figure 3.1: Private school participation rate age 6–10, 1998-2011



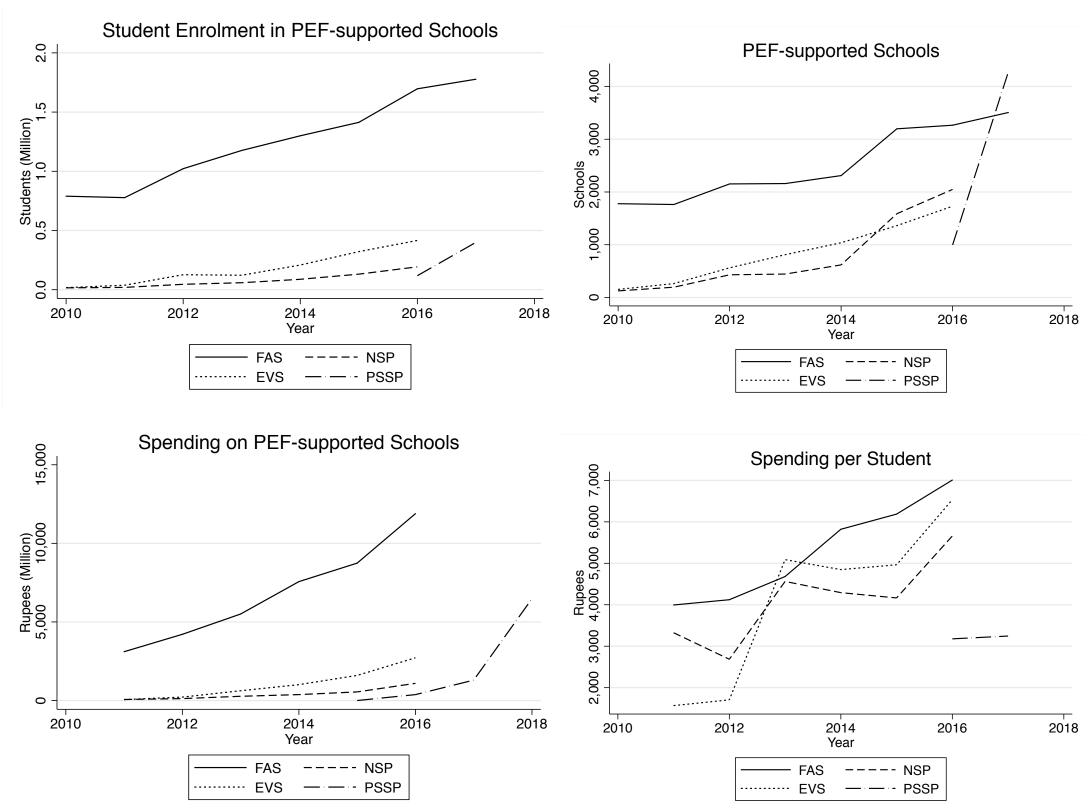
Source: (Nguyen & Raju, 2014)

Pakistan also has a long history of government engagement with the non-state sector in education. The federal education policy of 1972 declared "education will be made free and universal up to Class X [Ten] for all children throughout the country ... in both Government and privately-managed schools. Private schools will be suitably supported for the loss of fees incurred by them" (Bengali, 1999). This position was repeated in 1992 when a new policy "declared the State's intent for emphasizing the private sector's role in education through "viable partnership[s]" … and reiterated strongly in 2001: "Acknowledging the shift in government's role from being a provider to a facilitator […] it is vital to rethink the parameters of public private partnership in the provision of education" (Malik, Bari, Muzaffar, & Khan, 2015). More recently the 2017 National Education Policy has outlined specific objectives to "encourage, facilitate and regulate private sector education" and to "promote regulated and monitored Public-Private Partnership[s] for educational development." The policy also recommends that "Innovative programmes such as "Adopt a School" programme shall be continued" (Government of Pakistan Ministry of Federal Education and Professional Training, 2017).

Most Public Private Partnerships (PPPs) in Pakistan have been either subsidies for private schools to enable them to accept pupils at reduced cost, or vouchers provided to students to enable them to pay the fees at private schools. More recently provincial governments have begun to explore 'contract management' PPPs similar to US Charter schools, in which private organisations are contracted to take over the management and operations of existing public schools, which remain government owned, financed, and regulated.

Education PPPs in Punjab are managed by the Punjab Education Foundation (PEF), itself a quasi-independent body. PEF was established in 1991, and made autonomous in 2004. PEF has three main programmes of support for private schools, all of which have grown over the past decade. The largest programme is the Foundation Assisted Schools, through which 1.9 million children are educated in registered private schools, with fees paid by PEF. The Education Voucher Scheme supports 500,000 children. PEF also has a New School Programme, which has contracted private organisations to build around 2,000 schools in remote and under-served areas, which have enrolled 250,000 students.

Figure 3.2: Trends in PEF-supported Schools



Note: All data is from Punjab Education Foundation (PEF) Annual Reports. Real spending is calculated using World Bank Consumer Price Inflation. FAS is the acronym for "Foundation Assisted Schools". NSP is for "New Schools Program". EVS is for "Education Voucher Scheme". PSSP is for "Public School Support Programme".

Several papers have estimated the effects of subsidies and voucher programmes in Pakistan on test scores, generally finding positive effects on both enrolment and learning outcomes (Kim et al., 1999; Alderman et al., 2003; Barrera-Osorio et al., 2013; Amjad & MacLeod, 2014; Barrera-Osorio and Raju, 2015; Barrera-Osorio et al., 2017; Andrabi, Das, Khwaja, Ozyurt, & Singh, 2018). One descriptive study considers the “adopt-a-school” programme in Punjab and Sindh covering 1,500 schools (Malik et al 2015). A full summary is presented in Table 3.2. This is the first study to estimate the causal effect of the PSSP programme, which has transferred over ten percent of government primary schools in the province to private management in its first year.

Table 3.2: Literature on Public-Private Partnerships in Pakistan

Author	Date	Province	PPP Type	Outcome Type	Outcome	Study Type
Alderman et al	2003	Balochistan	Subsidy	Enrolment	Positive	RCT
Amjad & Macleod	2014	National	Subsidy & Voucher	Learning	Positive	OLS
Andrabi et al	2018	Punjab	Subsidy	Learning	Positive	RCT
Barrera-Osorio et al	2013	Sindh	Subsidy	Learning	Positive (0.16 SD)	RCT
Barrera-Osorio & Raju	2015	Punjab	Subsidy	Enrolment	Positive	RD
Kim et al	1999	Balochistan	Subsidy	Enrolment	Positive	RCT
Malik et al	2015	Punjab & Sindh	Contract Schools	Learning	Positive	PSM

Note: This table presents a summary of all identified studies on PPPs education in Pakistan

The Public School Support Programme (PSSP)

In December 2015 the Punjab government announced that it would transfer management of 5,000 failing government schools (12 percent of all primary schools) to private operators. The 'Public School Support Programme' (PSSP) began Phase 1 just four months later, at the start of the next school year in April 2016. Phase 2 began after the summer break in August 2016, and Phase 3 at the start of the following school year in April 2017. So far, 4,276 schools with over 500,000 students have been transferred. The Punjab Education Foundation (PEF) tendered schools competitively, with eligibility criteria laid out for two categories of bidders – organisations (existing school operators and non-governmental organisations (NGOs)), and individuals. Government received 19,000 applications for the first phase of 1,000 schools. Organisations were prioritised over individuals in the bidding process. Organisations with experience running schools were prioritized further. There are now around 2,600 schools contracted to organisations (each organisation has at least ten schools) and 1,700 schools contracted to individuals. Organisations are paid 700 Pakistani Rupees per child per month, and individual operators 550 Rupees[21]. This is equal to less than half of government per pupil spending in public schools (1,507 Rupees per pupil per month)[22]. Students may enter Katchi grade (Kindergarten) at age four. Each PSSP school has a 2-year contract, with renewal subject to adequate performance on 'Quality Assurance Tests' (QATs). In the initial phase PSSP schools fell under the remit of PEF, but have since been transferred to a separate

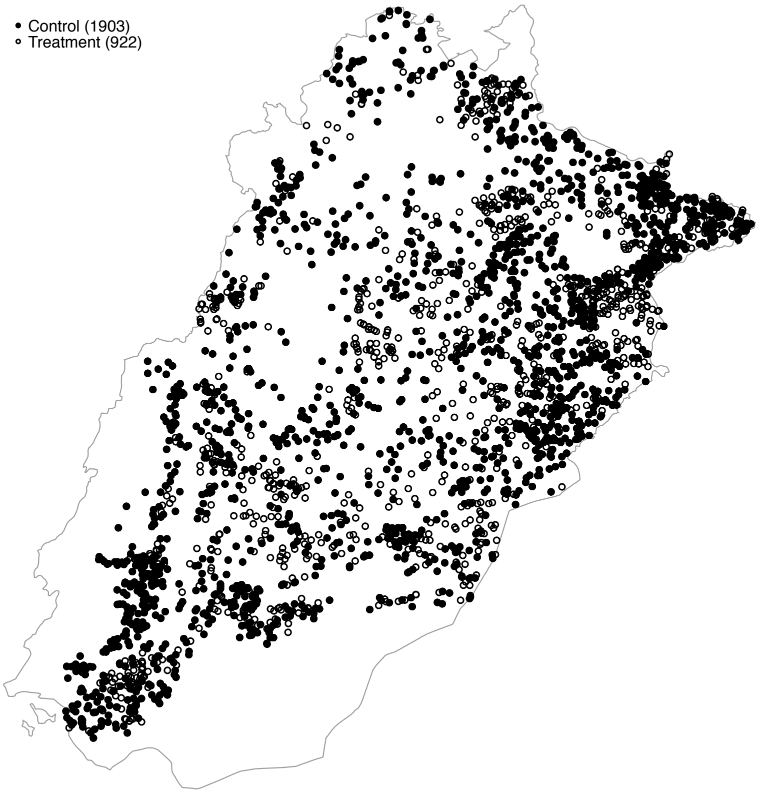
---

[21] 550 Rupees is the same amount provided to private schools through the Foundation Assisted Schools (FAS) programme (reaching 1.8 million students), and the Education Voucher Scheme (EVS) (reaching 0.5 million students).

[22] Institute of Social and Policy Sciences report on Public Financing of Education in Pakistan 2010-11 to 2016-17.

authority, the Punjab Education Initiative Management Authority (PEIMA). PSSP schools are spread across the province, as shown in Figure 3.3.

Figure 3.3: Map of Treatment (Phase 1) and Comparison (Phase 3) Schools



Note: This figure demonstrates that there is no geographical clustering of schools by phase. We refer here to Phase 1 schools as “treatment” and Phase 3 as “comparison”.

PSSP schools are not allowed to charge fees, make profit, or select their students, though anecdotally some schools may have engaged in some selective admissions (M. Afridi, 2018). PSSP schools can hire their own teachers and head teachers at market salaries (teacher salaries in private schools are typically less than half of those in government schools). Incumbent government teachers were transferred to other government schools. Private school teachers in Pakistan are typically young unmarried women with less formal education, whereas public school teachers are more likely to be older men with more formal education. Despite the difference in teacher qualifications, private schools are typically able to elicit greater effort from their teachers, and achieve comparable or better results (Andrabi et al., 2008). A small sample of PSSP schools reported in Afridi (2018) suggests that teachers in PSSP schools are similarly qualified to those in other PEF supported private schools. PSSP schools teach the regular curriculum, and their students sit both the standard Grade 5 exams administered to students in both public and private schools by the Punjab Examinations Commission (PEC), as well as the Quality Assurance Test (QAT) exams administered to PEF-partnering private schools for all grades. PSSP schools must meet minimum standards in the QAT for continued participation in the programme, and may be eligible for financial bonuses for good performance. The School Education Department (SED) maintains ownership of buildings and responsibility for maintenance. PSSP school operators are required to submit quarterly expenditure statements detailing how income from government was spent. Payments are made monthly to school bank accounts. For the first six months schools are due a fixed amount regardless of enrolment – after this period they are due a variable amount based on the number of enrolled students (regardless of grade). Schools continue to teach in the existing medium of instruction, using textbooks provided by the Schools Education Department (SED) or PEF. Students wear the same uniforms as those worn in government schools.

Schools were eligible to be selected if they met any one of five criteria - being a) overcrowded, b) under-utilised, c) with low enrolment, d) low exam pass rates, or e) entirely non-functional. Overall 10,664 schools met at least one of these criteria, leaving a large number of schools that may still be selected into any future Phase 4. Executive District Officers (EDOs) were responsible for selecting which schools to recommend for transfer. The selection criteria are consistent across the 3 phases, with one change being the expansion of the definition of poor learning outcomes from a 0% pass rate in Phase 1 to below 25% from Phase 2, as few schools met this criterion in Phase 1.

Table 3.3: PSSP Selection Criteria

Criteria	Definition	Schools selected (Phase 1 – 3)	All Eligible Schools in Punjab	% Selected
1. Multi grade & over crowded	1 teacher, needs more classrooms, & above 80 students	103	943	11
2. Under utilised	2 or more teachers & below 30 students	430	1,464	29
3. Low enrolment	Fewer than 21 students in Grade 1-5	1,504	5,320	28
4. Poor learning outcomes	0% PEC passing rate (Phase 1 handover) or Less than 25% passing rate (Phase 2 handover)	872	1,317	66
5. Non-functional & closed	Non-operational / Merged	520	1,620	32
Total		3,429	10,664	32

Note: Schools were eligible to be recommended for inclusion in the PSSP if they fulfilled at least one of the selection criteria. McKinsey (2017) reports aggregate numbers selected according to each criterion. Our data does not associate individual schools with specific criteria.

3.3 Methodology

Our key challenge in estimating the causal effect of school conversion is finding an appropriate counterfactual. As schools were selected for PSSP based on low enrolment and exam scores, they are not comparable to non-PSSP schools.

We use a standard difference-in-difference strategy similar to that used by Abdulkadiroğlu, Angrist, Hull, & Pathak (2016) and Eyles et al. (2017) in the context of US Charter school and UK Academy converters. We compare the change in outcomes for early converting schools with the change in outcomes for later converters. We assign schools converted in Phase 1 as treatment schools, and schools selected later in Phase 3 as comparison schools. The change in outcomes is then presented for the one school year when Phase 1 treatment schools had already been converted and Phase 3 comparison schools had not yet. Both treatment and comparison schools were selected into the programme according to the same criteria. Whether schools were selected into Phase 1 or Phase 3 was essentially arbitrary, as the programme was launched just four months after being announced. Figure 3.3 demonstrates that there is no geographical clustering of Phase 1 and Phase 3 schools, which are evenly spread across the Province.

We estimate the following standard difference-in-difference equation in which T is a binary indicator for treatment status, 'Post' is a binary indicator for pre or post status, $\boldsymbol{\gamma_i}$ are school fixed effects, δ_t are year fixed effects, and our main coefficient of interest is $\boldsymbol{\beta_3}$ looking at the effect of the interaction between treatment and post.

$$y_{it}=\alpha+\beta_1 T_{it}+\beta_2 Post_{it}+\beta_3 T_{it} Post_{it}+\gamma_i+\delta_t+\varepsilon_{it} \quad (2.1)$$

For $\boldsymbol{\beta_3}$ to be an unbiased estimator of the true causal effect, we need two assumptions. First, that there is no time-variant group-specific unobservables – something unobserved that is specific to treatment schools that also determines student outcomes. Second, we assume that T the effect of time on outcomes is the same for both groups. This second assumption is the parallel trends assumption. Whilst we cannot test whether trends would have been parallel in the absence of the treatment, we can test a corollary of the assumption – that pre-trends before treatment assignment should also be parallel.

3.4 **Data**

Enrolment

We use data on enrolment from three sources; the Government Annual School Census, monthly monitoring data, and data collected from the Annual Status of Education Report (ASER). The timing of data collection across each source and the stages of the PSSP reform are shown in Table 3.5.

The primary source is the annual Government of Punjab School Census data (also known as EMIS - Education Management Information System). Head teachers report in October each year[23]. The data includes student enrolment by

[23] For 2017 we use enrolment data from monthly independent monitors collected from schools.

grade, school facilities, staffing, and location. The main weakness of the EMIS data for our purpose is that the outcome variable of interest (enrolment) is self-reported by schools. We address this by crosschecking the data against other sources. We compile EMIS data for all schools from 2012-13 to 2017-18.

Monitoring and Evaluation Assistants (MEAs) from the Programme Monitoring and Implementation Unit (PMIU) of the School Education Department visit schools each month. Assistants observe student enrolment, teacher presence, and the availability of utilities. A team of 856 monitors covers 47,725 government schools (on average 56 schools each). Assistants enter data using a tablet-based mobile app, allowing for built-in validation checks. We only have data from this source for part of 2017, after the conversion of Phase schools, so we cannot use it for our main analysis, but can use this independently collected source of data as a robustness check on the self-reported EMIS data where there is overlap. Enrolment for PSSP schools reported by the independent monitors is on average 1.8 students lower than that self-reported by head teachers through the EMIS. The correlation between the two measures is 0.89.

The Annual Status of Education Report (ASER) is a citizen-led survey of students and schools. The survey covers over 250,000 children each year, tested at their home rather than at school to capture those not enrolled. The survey also gathers basic data about the government and private schools that are available to children, including school enrolment. The survey is conducted between September and November each year. The ASER dataset is limited to a sample of 734 schools in Punjab (of 37,078 total primary schools). There are only 24 PSSP Phase 1 or 3 schools in the ASER dataset. However the dataset does allow for the comparison of figures for the 734 Punjab schools that do appear in both datasets. Enrolment for PSSP schools reported in ASER is slightly higher than enrolment reported in the EMIS, though this difference is not statistically significant. The correlation between the two measures is 0.69.

Learning Data

For learning we use three different sources of data, from the Punjab Examinations Commission (PEC), the PMIU Learning and Numeracy Drive (LND), and the Punjab Education Foundation (PEF) Quality Assurance Tests (QAT).

The primary source is the Punjab Examinations Commission (PEC) data. All students in both government and private schools are tested at Grade 5 and Grade 8 if they want to progress to the next level of schooling. These exams are high stakes for the student but not for the school. Exams are sat in February each year. Students must be at least 8 years old to sit the Grade 5 exam. The exam is 90 - 150 minutes long for each subject, with 34 items (questions). 30 items are multiple-choice focused on knowledge, and four are open-ended focused on comprehension. PEC reports average percentage marks (from 1 – 100) for all exam candidates from each school in five subjects; Urdu, Mathematics, English, Science, and Islamiat. We standardise test scores by subject and year.

Second, we use Learning and Numeracy Drive (LND) collected on a monthly basis from all schools. School Monitoring Officers visit all public schools monthly and test a sample of five Grade 2 and five Grade 3 students using a tablet-based app. Each student is given seven randomly selected multiple choice question items in Urdu, English, or Mathematics. The majority (83 percent) of schools are tested at least five times in the seven months for which we have data.

Third, PEF collects test data from all schools under its jurisdiction as part of its accountability framework. All schools are private schools except for PSSP schools in its initial phases. At the primary school level, tests are conducted in two randomly selected classrooms. Students sit a two-hour exam covering four subjects; English, Urdu, Science, and Mathematics. Schools that fail two consecutive QATs lose their entitlement to public funds. For a school to pass, at least half of students must pass the exam. For students to pass they must get at least 40 percent of the available marks. In our data, we have the average pass rate for all students at each school. The majority of schools in the PSSP first phase are managed by non-governmental organisations. Of these schools, 43 percent passed the QAT Table 3.4. PEF provides schools with model papers and past exam papers to enable them to prepare students for the exams.

Table 3.4: QAT Pass Results for Phase 1 NGO Schools

	Schools	Percentage of schools passing
All Phase 1 NGO Schools	626	0.43
Ghazali Education Trust	30	0.90
Learning Zone	10	0.80
National Rural Support Programme (NRSP)	100	0.60
Idara-e-Taleem-o-Aagahi (ITA)	30	0.53
Akhuwat	100	0.41
Punjab Rural Support Programme (PRSP)	70	0.41
CARE Foundation	100	0.40
The Citizens Foundation (TCF)	80	0.36
Ghazali Society	45	0.24
Developments in Literacy (DIL)	31	0.19
Muslim Hands	30	0.13

Note: Pass rates for NGO operators in Phase 1 are reported in
https://tribune.com.pk/story/1485071/57-ngo-run-public-schools-fail-pef-assessments/

Phase 1 (Treatment) schools began operation at the start of the school year in April 2016. The primary outcome data we use comes from the following year EMIS collected in October 2016 (six months after conversion) and PEC exams from February 2017 (ten months after conversion). Phase 3 (Comparison) schools then began operations at the start of the school year in April 2017. We discard data on Phase 2 schools that began four months into the school year and so may have faced considerable disruption.

Table 3.5: Data and Programme Timing

Month	PSSP Programme	Enrolment Data			Learning Data		
		EMIS	MEA	ASER	PEC	LND	QAT
Oct-15		2015-16		Y			
Nov				Y			
Dec							
Jan-16							
Feb					G5		
Mar							
Apr	Phase 1 (Treatment)						
May							
Jun							
Jul							
Aug							
Sep				Y			
Oct		2016-17		Y			
Nov				Y			
Dec							
Jan-17			Y			G2-3	
Feb			Y		G5	G2-3	
Mar			Y			G2-3	P1 Endline
Apr	Phase 3 (Control)		Y			G2-3	P3 Baseline
May			Y			G2-3	
Jun							
Jul							
Aug			Y			G2-3	
Sep			Y	Y		G2-3	

Note: This table outlines the timing of the PSSP programme, and how this overlaps with the available data on enrolment and learning.

Although our identification does not rest on baseline balance in covariates between schools, we nonetheless present baseline descriptive statistics for the groups of schools. Differences between treatment and comparison schools are statistically significant but small for prior enrolment, years in operation, number of classrooms, and classes. Differences are not statistically significant for the number of books. PSSP schools are all primary schools. PSSP schools are smaller than average. Phase 1 schools had 77 students and Phase 3 schools 68 students, compared with 114 students for all other public primary schools.

Table 3.6: Baseline Balance (2015-16)

	Treatment (Phase 1)	Comparison (Phase 3)	Non PSSP Primary Schools	Diff (Phase 1 – Phase 3)	P – Value
N (Schools)	995	1,977	34,163	.	.
Enrolment (Students)	77	68	114	9	0.000
Teachers	2.3	2.1	3.2	0.2	0.000
PEC Pass Rate	0.65	0.64	.		0.38
PEC Average Marks	53.8	53.2	.	0.6	0.27
PEC Candidates	5.9	5.4	.	0.5	0.012
Years in operation	37	35	39	2	0.000
Classrooms	2.8	2.5	3.2	0.3	0.000
Classes	5.9	5.6	5.8	0.3	0.000

Note: This table presents characteristics of PSSP schools prior to conversion. We don't have data for PEC exams for non-PSSP schools.

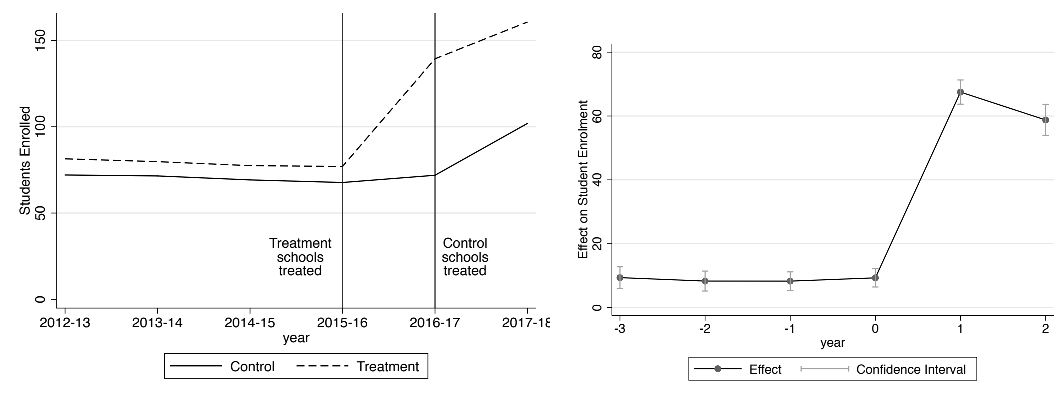
3.5 Results

In the following analysis, we first focus on enrolment as an outcome before moving to learning outcomes. For each outcome, we first present graphs of the average outcomes over time. Second, average outcomes pre- and post- reform. Third, the same analysis in an Ordinary Least Squares (OLS) regression framework including comparisons for school and time fixed effects.

Enrolment

The first outcome considered is the total number of enrolled students, as reported by the head teacher through the EMIS system. We observe a parallel trend in enrolment for treatment and comparison schools between 2012 and 2015, with a break between 2015-16 and 2016-17 when treatment schools are treated and their enrolment rises. In the 2017-18 school year, enrolment continues to rise in treatment schools. Comparison schools are also treated in 2017-18, and their enrolment begins to increases in parallel with treatment schools. The visual inspection of the parallel trends is confirmed by statistical test, which shows that the interaction between treatment status and the year indicator is insignificant in the pre-treatment period.

Figure 3.4: Enrolment trends in treatment & comparison schools



Note: The left panel presents trends in average student enrolment numbers for treatment (Phase 1) schools and comparison (Phase 3) schools. The right panel presents estimated treatment effects by year (controlling for lagged enrolment). Data for 2012-2016 is self-reported by head teachers in annual school census carried out in October. Independent monitors collect data for 2017 from the school register in August.

We next present average enrolment levels for the years immediately pre- and post- treatment. The difference in change in means across treatment and comparison is statistically significant. Enrolment in treatment schools increased by 58 students (60 percent) more than in comparison schools (Table 3.7).

Table 3.7: Average (mean) enrolment

Phase	Mean Enrolment 2015-16	Mean Enrolment 2016-17	Difference	P - Value	Schools
Comparison (Phase 3)	68	72	4		1,945
Treatment (Phase 1)	77	139	62		938
Difference	9	67	58	0.00	

Note: This table shows the simple differences in mean enrolment outcomes for comparison and treatment schools pre- and post- reform.

We then estimate the effect of treatment in a regression framework, including all previous years of data, year fixed effects (column 1), school controls (column 2), and school fixed effects (column 3). The coefficient is stable at 48 to 49 students across specifications (Table 3.8).[24] In studies with a large number of time periods, serial correlation in both the outcome and the independent variable of interest may be a concern, biasing our standard errors. Following Bertrand et al. (2004) we address this concern by collapsing the data into a single pre- and post- reform period, finding qualitatively similar results.

Looking at heterogeneity by school type, enrolment increases by more in single-sex schools than in mixed schools. We find no difference in effect size on enrolment by other prior school characteristics, including the criteria by which they were selected into the programme, their initial size, or their age[25]. Schools contracted to an NGO increased their enrolment by less than those contracted to individuals.

[24] We report consistent estimates based on independently collected data (from a very small sample) in Table A3.2.

[25] Our data does not include which school was selected according to which criteria, and so we reconstruct an estimate for each school of which criteria they were selected according to, based on the definition in Table 3.3. A comparison with actual aggregate numbers for each criterion and numbers based on our reconstruction are shown in Table A3.1.

Table 3.8: Effect of Treatment on Enrolment

	(1)	(2)	(3)	Boys	Girls	Mixed
Treatment x Post	48.607***	48.611***	48.366***	56.334***	53.468***	45.509***
	(1.488)	(1.499)	(1.484)	(3.645)	(4.201)	(1.764)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School Controls		Yes				
Baseline Comparison Group Mean	76.6	76.6	76.6	84.5	80.4	74.2
N	17,099	16,701	17,099	3,316	2,071	11,478
N (Schools)	2,884	2,794	2,884	555	346	1,921
R-squared	0.207	0.313	0.335	0.385	0.344	0.323

Note: The dependent variable is the number of students enrolled. Columns (1) – (3) show all schools, and (4) – (6) show boys, girls, and mixed schools, respectively. School controls include prior number of years in operation, number of classrooms and classes, and district fixed effects. Coefficients on the treatment and post dummies are omitted.
* p<0.1, ** p<0.05, *** p<0.01

Breaking down this change in enrolment by grade, we observe that the increase is concentrated in Katchi grade (Kindergarten) (40 students) and the early grades. The difference in Grade 5 is just one additional student in treatment over comparison schools.[26]

[26] We do not have data on student attendance, but unpublished analysis by McKinsey (2017) of MEA data suggests a slight decrease in the student attendance rate of Phase 1 schools after conversion.

Table 3.9: Effect on Enrolment by Grade

	K	1-5	1	2	3	4	5
Treatment x Post	39.928***	18.488***	6.961***	3.923***	3.956***	1.857***	1.027***
	(1.254)	(1.022)	(0.439)	(0.311)	(0.259)	(0.225)	(0.203)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Comparison Group Mean	26.4	40.9	12.3	9.8	7.4	6.5	4.9
N	5,588	5,588	5,588	5,588	5,588	5,588	5,588
N (Schools)	2,794	2,794	2,794	2,794	2,794	2,794	2,794
R-squared	0.505	0.290	0.207	0.174	0.167	0.069	0.044

Note: The dependent variable in each regression is the number of students in that grade. "K" stands for "Katchi" which is equivalent to Kindergarten. Coefficients on the treatment and post dummies are omitted. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Enrolment Spillovers

Were newly enrolled children previously out of school, or enrolled elsewhere? While we find a large effect of treatment on enrolment at PSSP schools, an important policy question is whether total enrolment increased. Did PSSP schools attract students who would not otherwise have attended school at all, or who were already enrolled in other schools. The data presented so far on enrolment at the school level does not allow us to say where new students came from. That most of the new enrolment is in Katchi grade rather than higher grades does suggest that many new students were not yet in school. But we don't know how many of those would have otherwise enrolled elsewhere. Schools may also have more actively recruited younger children, as government holds schools accountable for QAT test results in grades two to five.

We perform three other tests. First, we focus only on those villages in which the PSSP school is the only public school. Enrolment in these schools is less likely to be affected by spillovers from other nearby schools. There are 23 union councils with only a treatment PSSP school, and 60 with only a comparison PSSP school. The estimated effect on enrolment in this sub-sample is 71 students – larger than in the full sample.

Second, looking at enrolment at the closest neighbouring public school should directly capture any spillover[27]. We link all treatment and comparison schools to their nearest neighbouring school. The median nearest neighbour is within 0.5km of the school, and 75 percent of neighbours are within one kilometre. We then estimate the same difference-in-difference model as in the previous section, but replace the outcome as the enrolment level at the nearest neighbour rather than the treatment school itself. Results suggest that schools who's nearest neighbour is a PSSP school also see an increase in enrolment. This is the opposite of the negative effect that we would expect if enrolment in treatment schools were driven by the recruitment of students from neighbouring schools. This result is robust to dropping neighbours further than one kilometre. As discussed below under mechanisms for the main effect, new PSSP school operators carried out community engagement in order to encourage parents to enrol their children. It is possible that this engagement led to positive spillovers for non-PSSP public schools also. We see no difference in the effect for small schools or for schools operated by NGOs rather than individuals.

[27] We do not have data on enrolment in private schools

Table 3.10: Effect of Neighbouring a Treatment School on Enrolment

	(1)	(2)	(3)	(4)
	All	All	K	All Primary
Neighbour Treated X Post	-0.790	11.504***	14.337***	1.649
	(7.878)	(3.903)	(3.368)	(2.125)
Year FE	Yes	Yes	Yes	
School FE		Yes		
Baseline Comparison Group Mean	124.2	124.2	26.4	40.9
N	8,956	8,956	824	824
N (Schools)	2,589	2,589	414	414
R-squared	0.015	0.119	0.234	0.132

Note: The dependent variable is enrolment at the closest neighbouring school to each PSSP treatment school. "K" stands for "Katchi" which is equivalent to Kindergarten. Coefficients on the treatment and post dummies are omitted. * p<0.1, ** p<0.05, *** p<0.01

Third, we aggregate the data to the Union Council level. There are 6,646 union councils in Punjab. The median Union council has 5 schools. For each union council, we calculate total enrolment by year, and the proportion of all schools that were Phase 1 (treated) schools. We then report the same difference-in-difference analysis. In this case the treatment variable is the share of schools in the council that were in Phase 1 (treated). Column (1) in Table 3.11 shows the effect for the full sample of moving from zero treated schools to all treated schools. This estimate is biased as we are no longer comparing to later converting schools but to all schools, and many union councils with no PSSP school at all. Adding union council fixed effects reverses the sign on the diff-in-diff coefficient. Finally, we restrict the sample in columns (3) and (4) to only those union councils in which there was at least one PSSP school. These estimates are consistent with and without union council fixed effects.

Table 3.11: Effect on Enrolment at Union Council Level

	(1)	(2)	(3)	(4)
Treatment X Post	-22.273*	34.951***	47.050***	46.149***
	(12.041)	(7.791)	(10.465)	(9.889)
Year FE	Yes	Yes	Yes	Yes
Union Council FE		Yes		Yes
Baseline Mean	255	255	162	162
Treated share	0.017	0.017	0.161	0.161
N	34,607	34,607	4,187	4,187
N (Units)	6,645	6,645	707	707
R-squared	0.020	0.072	0.027	0.075

Note: The dependent variable is total enrolment in the district council. Treatment is defined as the share of schools in the union council with a Phase 1 treated PSSP school. Coefficients on the treatment and post dummies are omitted. * p<0.1, ** p<0.05, *** p<0.01

Mechanisms

How and why do schools increase student enrolment? First, schools have a strong incentive to maximize enrolment as they are funded on a per student basis. Parents face no change in price that remains zero, so schools are forced to compete on quality (or at least indicators of quality that are visible to parents). One clearly visible indicator of quality is class sizes. We observe large increases in the number of teachers (from two to five), reducing pupil-teacher ratios from 35 to 25, despite the concurrent increase in student enrolment. This increase in teachers is enabled by the substantially lower market rate for salaries in private schools - 1,407 rupees ($12) per month, compared to 7,671 rupees ($66) in government schools (Bau & Das, 2017). Analysis of MEA data by (McKinsey, 2017) found no change in teacher presence rates for PSSP Phase 1 and 2 schools after conversion, and no change in school facilities (boundary walls, toilets, electricity, drinking water).

Table 3.12: Effect of Treatment on Class Size (Pupil-Teacher Ratio)

	(1)	(2)	(3)	(4)	(5)	(6)
		Pupil teacher ratio			Teachers	
Treatment x Post	-9.323***	-8.816***	-9.324***	3.068***	3.062***	3.087***
	(1.052)	(1.066)	(1.063)	(0.084)	(0.085)	(0.085)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School Controls		Yes			Yes	
School FE			Yes			Yes
Baseline Comparison Group Mean	34.5	34.5	34.5	2.1	2.1	2.1
N	5,664	5,491	5,664	5,679	5,504	5,679
N (Schools)	2,879	2,791	2,879	2,880	2,791	2,880
R-squared	0.013	0.135	0.038	0.478	0.534	0.551

Note: The dependent variable in columns 1-3 is the pupil-teacher ratio, and in columns 4-6 is the number of teachers. School controls include prior number of years in operation, number of classrooms and classes, and district fixed effects. Coefficients on the treatment and post dummies are omitted. * p<0.1, ** p<0.05, *** p<0.01

Second, there are some reports that some school operators passed on some of their subsidy directly to parents. One provider the Punjab Rural Support Programme is reported to have offered parents 500 rupees a month per child.[28] Other providers conducted weekly parent teacher meetings and events to involve parents in the school, or involved the local mosque leader to generate community buy-in. Third, some schools switched from being girl-only schools to being mixed, increasing their market size (M. Afridi, 2018).

[28] https://herald.dawn.com/news/1153868

Student performance

How do schools perform on learning? We estimate the same difference-in-difference model as before, now with the province-wide standardised Grade 5 exam test score data as the outcome. The outcome is the school-level average score. We first look at the number of students taking the exam. The number of students taking the exam in treatment schools increased from 5.8 in 2015-16 to 7.9 in 2016-17, an increase of two students. This difference falls to 1.5 students when controlling for school fixed effects (Table 3.13).

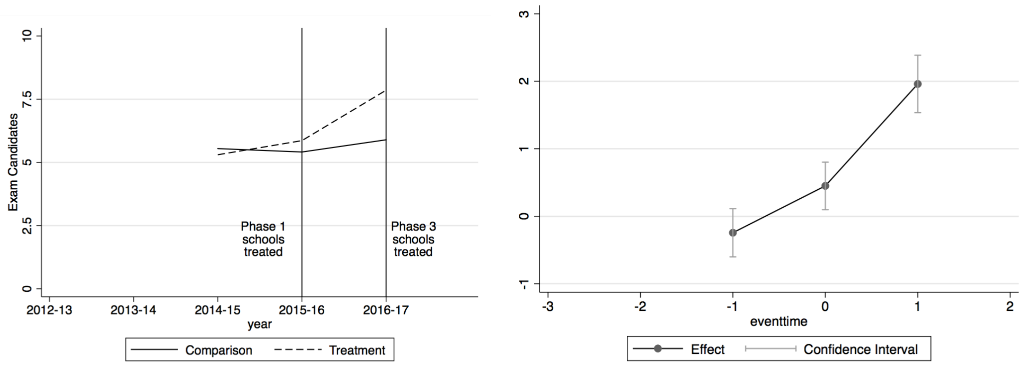
Table 3.13: Effect of Treatment on Grade 5 Exam Candidates

	(1)	(2)	(3)
Treatment x Post	1.960***	1.717***	1.513***
	(0.195)	(0.196)	(0.193)
Year FE	Yes	Yes	Yes
School Controls		Yes	
School FE			Yes
Baseline Comparison Group Mean	5.5	5.5	5.5
N	7,592	7,472	7,592
N (Schools)	2,780	2,726	2,780
R-squared	0.023	0.214	0.071

Note: The dependent variable is the number of Grade 5 examination candidates. School controls include prior number of years in operation, number of classrooms and classes, and district fixed effects. Coefficients on the treatment and post dummies are omitted.
* p<0.1, ** p<0.05, *** p<0.01

The prior trend in the average number of exam candidates is close to parallel in treatment and comparison schools (Figure 3.5).

Figure 3.5: Trends in Grade 5 exam candidates

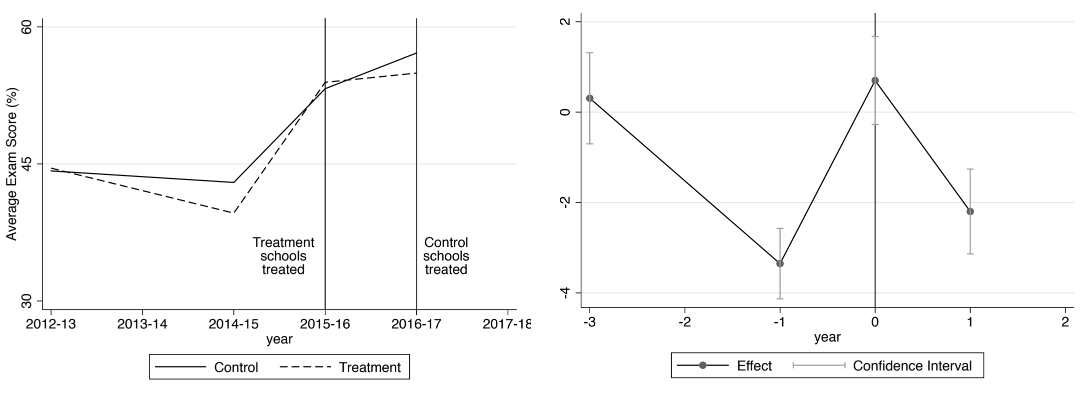


Note: The left panel presents trends in the average number of Grade 5 exam candidates for treatment (Phase 1) schools and comparison (Phase 3) schools. The right panel presents estimated treatment effects by year.

This increase in students taking the Grade 5 exam makes it difficult to interpret the effect on average test scores. We are unable to distinguish between a treatment effect on pre-enrolled students and a compositional effect from the entry of new candidates. The increase in new exam takers might come from marginal students with lower than average expected results.

The graph of average exam scores shows a slight relative decline for treatment schools after conversion. But prior trends are not parallel, casting doubt on the required parallel trends assumption for us to interpret these effects as causal.

Figure 3.6: Trends in Grade 5 Scores



Note: The left panel presents trends in average student exam scores for treatment (Phase 1) schools and comparison (Phase 3 schools). The right panel shows estimated treatment effects by year. All data is at the school-level from the Punjab Examinations Commission.

In the OLS regression framework we estimate a small negative effect on average test scores (-0.08 school-level standard deviations) and on Maths and English, (-0.1 school-level standard deviations). There is no effect on Urdu, Science, or Islam scores.[29]

[29] Note that the interpretation of effect sizes in terms of school average test score standard deviations is different to the interpretation of effect sizes in terms of individual student standard deviations. The variation in school average test scores is roughly half of the variation in student test scores, so to compare this effect size with estimates from impact evaluations using individual student data, one should divide this estimate by half.

Table 3.14: Effect on Grade 5 Test Scores, by Subject

	All	Urdu	Maths	Eng	Sci	Isl
Treatment x Post	-0.082**	-0.027	-0.105***	-0.105***	-0.032	-0.046
	(0.038)	(0.041)	(0.036)	(0.040)	(0.041)	(0.034)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE		Yes		Yes		Yes
Baseline Comparison Group Mean	0.0	0.0	0.0	0.0	0.0	0.0
N	9,758	9,758	9,758	9,758	9,758	9,758
N (Schools)	2,749	2,749	2,749	2,749	2,749	2,749
R-squared	0.328	0.095	0.409	0.231	0.278	0.174

The dependent variable is standardised to mean zero and standard deviation of one by year and subject. Coefficients on the treatment and post dummies are omitted.
* p<0.1, ** p<0.05, *** p<0.01

Ideally we would address the conflation of treatment and compositional effects by limiting our analysis to only students who would have taken the test without treatment. Unfortunately with our data we are unable to identify these students. What we can identify is the change in the total number of candidates entered by the school. We estimate heterogeneous effects by splitting the sample into schools that had more candidates, schools that had the same number, and schools that had fewer candidates than before. Here we see an important difference. Only schools that entered more candidates than before saw a decrease in test scores. Other schools saw no change in test scores. The number of candidates entered is endogenous to test scores and so we cannot be sure that these effects are causal, but the data is consistent with any negative effect being driven by a compositional effect rather than a treatment effect.

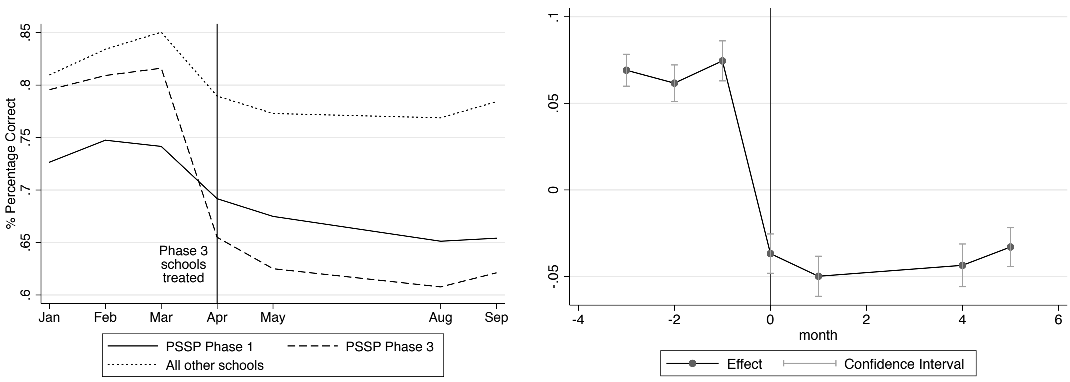
Table 3.15: Effect of Treatment on Test Scores, by School Type

Panel A: Schools with fewer candidates after treatment						
	All	Urdu	Maths	Eng	Sci	Isl
Treatment x Post	0.018	0.080	-0.035	-0.060	0.131*	-0.025
	(0.072)	(0.076)	(0.066)	(0.074)	(0.075)	(0.068)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE		Yes		Yes		Yes
N	3,114	3,114	3,114	3,114	3,114	3,114
N (Schools)	895	895	895	895	895	895
R-squared	0.314	0.095	0.410	0.210	0.262	0.169
Panel B: Schools with the same number of candidates after treatment						
	All	Urdu	Maths	Eng	Sci	Isl
Treatment x Post	-0.002	0.060	-0.060	0.042	-0.035	-0.004
	(0.092)	(0.099)	(0.100)	(0.100)	(0.113)	(0.086)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE		Yes		Yes		Yes
N	1,401	1,401	1,401	1,401	1,401	1,401
N (Schools)	425	425	425	425	425	425
R-squared	0.378	0.129	0.450	0.263	0.295	0.219
Panel C: Schools with more candidates after treatment						
	All	Urdu	Maths	Eng	Sci	Isl
Treatment x Post	-0.158***	-0.107**	-0.154***	-0.163***	-0.130**	-0.064
	(0.049)	(0.054)	(0.048)	(0.053)	(0.053)	(0.043)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE		Yes		Yes		Yes
N	5,243	5,243	5,243	5,243	5,243	5,243
N (Schools)	1,429	1,429	1,429	1,429	1,429	1,429
R-squared	0.327	0.090	0.401	0.239	0.287	0.168

Note: The dependent variable is standardised to mean zero and standard deviation of one by year and subject. Coefficients on the treatment and post dummies are omitted.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$

We can also use the monthly Literacy and Numeracy Drive (LND) data for 2017 to look at the change in performance for Phase 3 schools following conversion. Here we don't have data from before conversion of Phase 1 schools, so don't have the same natural treatment and comparison group as before. We do though have detailed student data on a monthly basis for all schools (this is a panel of schools and not students, who are randomly sampled in each round). Instead we show a comparison of Phase 3 schools before and after conversion with trends in Phase 1 schools (already converted) and all other non-PSSP schools. Here there is a large fall in test scores at Phase 3 schools following conversion. Again we are unable to distinguish between the treatment effect on pre-existing students and the compositional effect of the new students. Here the speed of the drop in test scores (in a single month) also seems to be more consistent with a compositional effect than a treatment effect. If teaching at converted schools were worse, we would expect to see a gradual divergence in test scores over time rather than a sudden drop. The LND data also tests different grades (2 and 3) to the earlier PEC data (grade 5). That the fall in test scores is larger in grade 2/3 than grade 5 is again consistent with a compositional effect, as the increase in enrolment was larger in lower grades. Enrolment increased by four students in grades 2 and 3, and by just one student in grade 5.

Figure 3.7: Trends in Grade 2-3 Test Scores (2017)



Note: The left panel presents trends in average student test scores for PSSP and all schools. Phase 1 schools were converted in the prior year, and Phase 3 schools in April. The right panel presents estimated treatment effects by year, in this case reversing the two groups and treating the Phase 3 schools (newly converting) as the treatment group and Phase 1 schools (already converted) as the comparison group.

The visual inspection of (Figure 3.7) is supported by the simple difference in average test scores (Table 3.16) and the OLS estimate with subject, grade, and school fixed effects (Table 3.17), which both suggest -0.11 percentage points lower test scores.

Table 3.16: Pre- and Post- Grade 2-3 Test Scores

	Jan-Mar 2017	Apr-Sep 2017	Difference
Phase 1	0.74	0.67	-0.07
Phase 3	0.81	0.63	-0.18
Other	0.83	0.78	-0.05
Phase 3 - Phase 1	0.07	-0.04	-0.11

Note: This table shows the simple differences in mean enrolment outcomes for comparison and treatment schools pre- and post- reform.

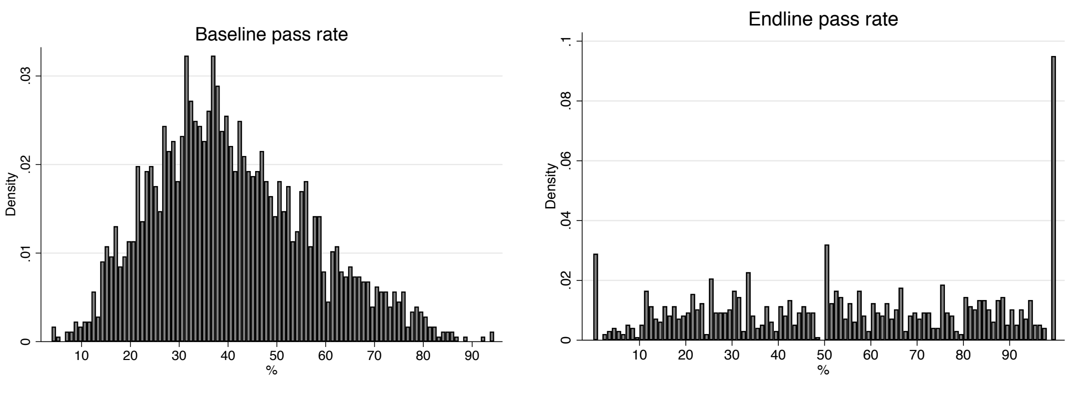
Table 3.17: Effect of Treatment on Grade 2-3 Test Scores

	(1)	(2)
Treatment x Post	-0.107***	-0.112***
	(0.006)	(0.006)
Subject FE	Yes	Yes
Grade FE	Yes	Yes
School Controls	Yes	Yes
School FE		Yes
N	70,967	70,967
N (Schools)	2,801	2,801
R-squared	0.103	0.102

Note: The dependent variable is standardised to mean zero and standard deviation of one by year and subject. Coefficients on the treatment and post dummies are omitted.
* p<0.1, ** p<0.05, *** p<0.01

In contrast to findings so far, results from Quality Assurance Tests (which are high stakes for schools) suggest improvement in treatment schools. Here we only have tests from a single point in time, but can compare results for Phase 1 schools at the end of their first year with Phase 3 schools at the start of their first year. Data is for the pass rate at the school. The average pass rate after treatment was 55 percent, compared with 41 percent before treatment. For schools to pass the QAT test, at least 50 percent of students must pass the test. 425 Phase 1 schools (43 percent) failed the QAT test at the end of their first year. We see some clear evidence of manipulation by schools. 31 schools report a pass rate of exactly 50 percent of students (the threshold required for the school to pass). Zero schools report a pass rate of 49 percent of students.

Figure 3.8: Quality Assurance Tests (QAT) (2017)



Note: This figure presents the frequency of schools reporting each student pass rate on the QAT test, with baseline pass rates on the left and endline pass rates on the right.

Bounds on the compositional effect

Arithmetically, we can estimate a bound on the size of any possible negative compositional bias. Test score data comes from Grades 2, 3, and 5. Although enrolment gains were concentrated in Katchi grade, there was a smaller increase in other grades as well. As numbers were so low to begin with, this means that 23 to 31 percent of students taking tests after the reform were new students. If we assume that the true treatment effect on the original students was zero, we can calculate how much worse the performance must have been amongst new students for the overall estimated average effect to be the size that it was. This calculation is laid out in Table 3.18 below. The implied performance of new students ranges from -0.35 standard deviations worse in Grade 5 to -0.45 in Grade 2.

Table 3.18: Minimum required bias from compositional effect consistent with a null treatment effect

	(1)	(2)	(3)	(4)	(5)
	Increase in candidates	Total new candidates	New Students as % of Total Students	Overall test score reduction (SD)	Minimum required negative effect of new students (SD)
Grade 5	1.5	6.5	0.23	-0.08	-0.35
Grade 3	4	13	0.31	-0.112	-0.36
Grade 2	4	16	0.25	-0.112	-0.45

Note: Column 4 shows the observed reduction in test scores for each grade. Column 5 calculates what the required relative average performance of the new students would have to be if they were entirely responsible for the overall reduction observed in column 4, if the true treatment effect on existing students were to be zero.

There could also of course have been selection in the other direction. Some new students may have come from private schools, and be more advantaged than the average student. We aren't able to place a bound on this possible positive selection.

3.5.1 Results for Different School Operators

We have data from six of the eleven NGO operators from Phase 1 that allow us to identify their schools in the data. Here we present results disaggregated by school operator for enrolment and pupil-teacher ratios. We omit analysis of learning given the uncertainty about the overall results. Each of the six operators has a positive effect on enrolment, though a smaller positive effect than other (anonymous) operators. All of the operators also increase the number of teachers hired. Two operators increase their teachers by more than average – TCF (5 teachers) and CARE (4 teachers). This leads to large reductions in pupil-teacher ratios in these schools by around 20 students (Table 3.19). Overall the NGOs have lower enrolment growth than other schools but a similar increase in teachers, leading to lower pupil-teacher ratios.

Table 3.19: Treatment Effects by Operator

	Enrolment	Enrolment	Teachers	Teachers	PTR	PTR
CARE	31.375***		3.950***		-19.619***	
	(3.091)		(0.202)		(2.230)	
DIL	13.815***		1.207***		-8.728***	
	(3.617)		(0.215)		(1.964)	
ITA	35.602***		2.865***		-16.357***	
	(4.294)		(0.295)		(4.565)	
Muslim Hands	45.932***		1.431***		-1.286	
	(5.606)		(0.292)		(4.467)	
NRSP	20.281***		1.166***		-9.140***	
	(2.861)		(0.127)		(3.178)	
TCF	38.354***		4.909***		-18.938***	
	(3.549)		(0.216)		(2.603)	
Other PSSP	59.439***		3.238***		-6.348***	
	(1.949)		(0.114)		(1.391)	
All NGOs		29.606***		2.812***		-13.687***
		(1.658)		(0.123)		(1.419)
Other PSSP		59.439***		3.238***		-6.348***
		(1.948)		(0.114)		(1.390)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Comparison Group Mean	76.2	76.2	2.1	2.1	34.2	34.2
N	17,068	17,068	5,548	5,548	5,529	5,529
N (Schools)	2,876	2,876	2,829	2,829	2,829	2,829
R-squared	0.346	0.345	0.610	0.554	0.051	0.043

Note: Coefficients on the treatment and post dummies are omitted. CARE is the “CARE Foundation”, DIL is “Developments in Literacy”, ITA is “Idara-e-Taleem-o-Aagahi”, NRSP is the “National Rural Support Programme”, TCF is “The Citizen’s Foundation”.

* p<0.1, ** p<0.05, *** p<0.01

3.6 Conclusion

In this paper we have estimated the effect of the fastest ever programme of contracting out government schools to private management. We find a large increase in enrolment and a modest decline in test scores. We are unable to say with full confidence how much of the increase in enrolment came from students who were not already in another school, or whether the observed decline in test scores is due to a negative treatment effect or a purely compositional effect. We also do not necessarily expect to see impacts on test scores in the relatively short time period for which we have data.

We are also unable to determine whether the effect on school enrolment is inherent to private management, or simply a function of a system of school financing in which schools are reimbursed on a per student basis.

Future research could usefully identify and track students who were enrolled in schools before transition, in order to get a clearer estimate of the actual treatment effect on learning outcomes.

Given the size of the non-state school market in Pakistan, it seems likely that there could be sufficient operator supply response if the government chose to further scale-up the programme.

PSSP schools reduce annual per pupil spending by government from 1,507 rupees to 550 rupees. Whilst this is a large margin, the total saving for 400,000 students is only 382 million rupees ($3.6m USD), or around 0.14 percent of the total provincial education budget (296 billion rupees or $2.5bn USD). This small potential saving has not actually been realized as existing teachers from PSSP schools were moved to other government schools rather than laid off. Hence an important question is how effectively those teachers are used in other schools and whether they fill gaps or duplicate existing effort. A further caveat is whether the payment of market salaries in PSSP schools is sustainable, or whether teachers may manage to lobby to receive regular government teacher salaries, as contract teaching assistants in Kenya (Sandefur, 2013) and India[30] have done.

[30] https://scroll.in/article/846589/in-uttar-pradeshs-botched-effort-at-regularising-contract-teachers-a-lesson-for-other-states

Appendix A3.1: Tables

Table A3.1: PSSP Selection Criteria

	Actual		Estimated			
Criteria	Schools selected Phase 1 – 3 (Percent)	Eligible Schools	Phase 1	Phase 2	Phase 3	Phase 1 - 3
1. Multi grade & over crowded	103	943	25	26	52	103
2. Under utilised	430	1,464	24	211	154	389
3. Low enrolment	1,504	5,320	73	527	455	1,055
4. Poor learning outcomes	872	1,317		397	1,285	1,682
5. Non-functional & closed	520	1,620	15	22		37
Total	3,429	10,664	137	1,183	1,946	3,266

Note: Schools were eligible to be recommended for inclusion in the PSSP if they fulfilled at least one of the selection criteria. Actual aggregate numbers selected according to each criterion are reported in columns 1 and 2. Our data does not associate individual schools with specific criteria, so we estimate school's selection criteria based on their enrolment, teachers, and test scores in 2015-16. Our estimate performs well for criteria 1 to 3, but over-estimates schools selected on criteria 4 and underestimates schools selected on criteria 5. Criteria 4 (poor learning outcomes) was 0% for Phase 1 and then increased to 25% for Phase 2 due to low numbers of schools with a 0% pass rate.

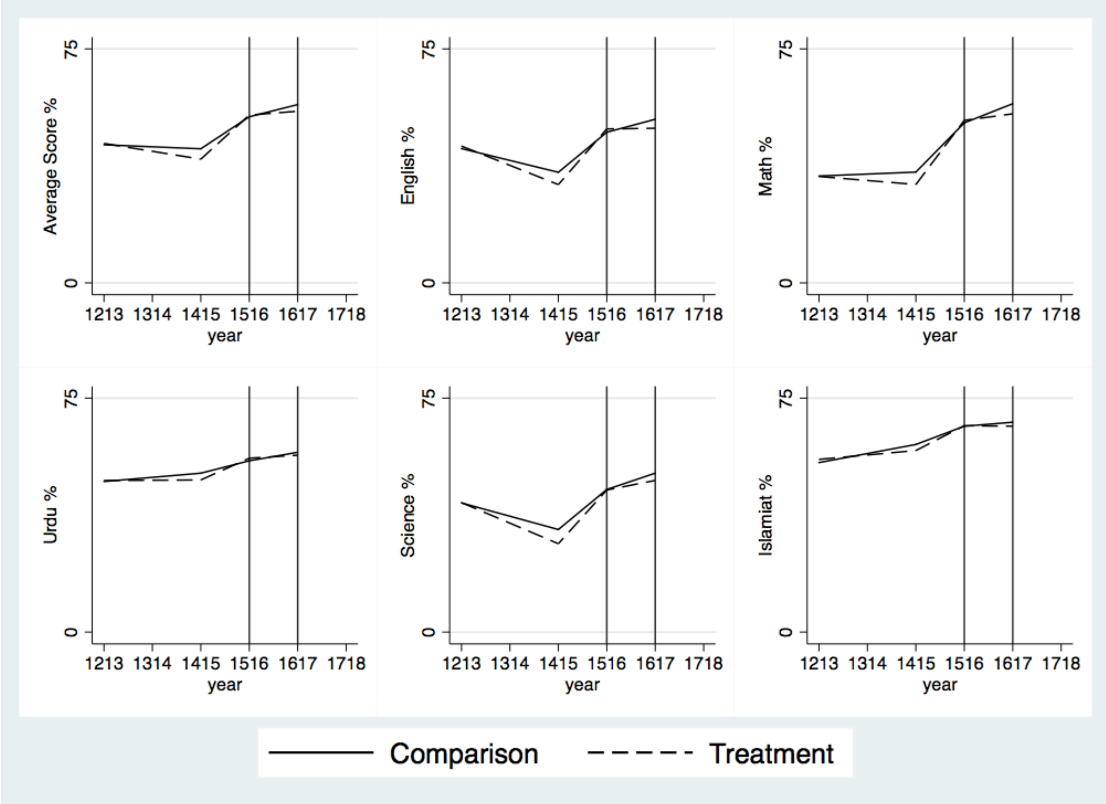
Table A3.2: Effect of Treatment on Enrolment

	EMIS Data	ASER Data
	(1)	(2)
Treatment x Post	48.297***	37.971
	(1.484)	(29.289)
Year FE	Yes	
Baseline Comparison Group Mean		
N	17,494	39
N (Schools)	2,972	36
R-squared	0.337	0.064

Note: The dependent variable in each column is total student enrolment. EMIS data is self-reported by head teachers. This is the main analysis reported in Table 3.8. ASER data is collected by an independent NGO from individual class registers. This data is available for only a small sample and so the results are not statistically significant, but the coefficient is similar to that using self-reported data.

Coefficients on the treatment and post dummies are omitted. * p<0.1, ** p<0.05, *** p<0.01

Figure A3.1: Trends in Grade 5 Test Scores by Subject



Conclusion

This thesis is comprised of three essays on the economics of education in developing countries, focused on the empirical micro-econometric analysis of public financing of schools operated by non-state actors.

Chapter One evaluates a policy on publicly subsidised private secondary schools in Uganda. Student value-added is higher in these schools than in government schools. The chapter explores the role of the quality of school management as a mediating factor in the performance differential, finding that private schools are on average no better managed than government schools, with the exception of those managed by an international charity. Future research could usefully address the question of how to improve school management at scale, and the role that performance management systems and school inspections can play.

Chapter Two evaluates a five-year private school voucher lottery programme in Delhi. This lottery was designed as a test of India's national Right to Education Act Section 12(1)(c), which reserves 25 percent of places at private schools nationwide for students from "economically weaker sections", with funding coming from government. Students who won the lottery and attended low-cost private schools performed slightly worse in Hindi and no different in Maths and English, or on various non-cognitive skills. Given our results consistent with imperfect information, a plausible role for government could be providing the public good of better information about school quality, as has shown to be effective at improving private school quality in the US (Hastings and Weinstein, 2008), India (Afridi et al., 2018) and Pakistan (Andrabi et al., 2017), or focusing on private school market failures in other areas (Andrabi et al., 2015).

Chapter Three evaluates a large-scale contracting out of public schools to private management in Punjab, Pakistan. Using a difference-in-difference framework, I estimate that failing government schools that are contracted out to private operators dramatically increase their enrolment, but that the effect on student learning is ambiguous. Future research could usefully identify and track students who were enrolled in schools before transition, in order to get a clearer estimate of the actual treatment effect on learning outcomes.

Overall these three studies highlight the variability in forms of public financing for independently operated schools, and the variability in quality. In all three policies, financing for non-state schools costs significantly less than equivalent spending in government schools. The chapters contribute to the literatures on public private partnerships in education, on the role of school management in determining student performance, and on school choice.

What is clear is that private schools are no panacea, but that they may play a role in providing more access to schooling, in some cases of somewhat better quality, and typically at lower overall cost. An important open question remains how best to regulate and oversee the quality of both private and public schools. One strength of public-private partnerships is that they demand that some kind of explicit contract is written down between government and school operators. In principle there is always some kind of implicit contract between government and schools whether they are private or public, but too rarely are priorities made explicit.

References

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots. *The Quarterly Journal of Economics*, *126*(2), 699–748. https://doi.org/10/dhnqpq

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., & Pathak, P. A. (2016). Charters without Lotteries: Testing Takeovers in New Orleans and Boston. *American Economic Review*, *106*(7), 1878–1920. https://doi.org/10/f8vsjd

Abdulkadiroğlu, A., Pathak, P. A., & Walters, C. (2018). Free to Choose: Can School Choice Reduce Student Achievement? *American Economic Journal: Applied Economics*, *10(1): 175–206*, 41.

Afridi, F., Barooah, B., & Somanathan, R. (2018). Improving learning outcomes through information provision: Experimental evidence from Indian villages. *Journal of Development Economics*. https://doi.org/10/gftrx8

Afridi, M. (2018). *Equity and Quality in an Education Public-Private Partnership: A study of the World Bank-supported PPP in Punjab, Pakistan*. https://doi.org/10.21201/2018.3002

Alderman, H., Kim, J., & Orazem, P. F. (2003). Design, evaluation, and sustainability of private schools for the poor: The Pakistan urban and rural fellowship school experiments. *Economics of Education Review*. https://doi.org/10/d2xf8d

Allende, C., Gallego, F., & Neilson, C. (2019). *Approximating the Equilibrium Effects of Informed School Choice* (p. 51). Princeton Industrial Relations Section Working Paper.

Altonji, J., Elder, T., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, *113*(1), 151–184. https://doi.org/10/c3xd4j

Amjad, R., & MacLeod, G. (2014). Academic effectiveness of private, public and private–public partnership schools in Pakistan. *International Journal of Educational Development*, *37*, 22–31. https://doi.org/10/f57wfc

Andrabi, T., Das, J., Ijaz Khwaja, A., & Zajonc, T. (2011). Do Value-Added Estimates Add Value? Accounting for Learning Dynamics. *American Economic Journal: Applied Economics*, *3*(3), 29–54. https://doi.org/10.1257/app.3.3.29

Andrabi, T., Das, J., & Khwaja, A. I. (2008). A dime a day: The possibilities and limits of private schooling in Pakistan. *Comparative Education Review*, *52*(3), 329–355. https://doi.org/10/dw8j9t

Andrabi, T., Das, J., & Khwaja, A. I. (2015). *Delivering Education: A Pragmatic Framework for Improving Education in Low-Income Countries*. Retrieved from http://www.hks.harvard.edu/fs/akhwaja/papers/deliveringeducation.pdf

Andrabi, T., Das, J., & Khwaja, A. I. (2017). Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets. *American Economic Review*, *107*(6), 1535–1563. https://doi.org/10/gbhzb4

Andrabi, T., Das, J., Khwaja, A. I., Ozyurt, S., & Singh, N. (2018). *Upping the ante: The equilibrium effects of unconditional grants to private schools* (No. WPS8563; pp. 1–96). Retrieved from The World Bank website: http://documents.worldbank.org/curated/en/489361534875600698/Upping-the-ante-the-equilibrium-effects-of-unconditional-grants-to-private-schools

Angrist, J., Bettinger, E., & Kremer, M. (2006). Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in

Colombia. *American Economic Review*, *96*(3), 847–862. https://doi.org/10/b8mmpx

Angrist, J. D., Bettinger, E., Bloom, E, King, E. M, & Kremer, M. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review, 92, 1535–59.*

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics*. https://doi.org/10/gftrqb

Asadullah, M. N. (2009). Returns to private and public education in Bangladesh and Pakistan: A comparative analysis. *Journal of Asian Economics*, *20*(1), 77–86. https://doi.org/10/fkdtqp

ASER Pakistan. (2017). *Annual Status of Education Report (ASER) 2016*. Retrieved from http://aserpakistan.org/document/aser/2016/reports/national/Annual-Status-of-Education-Report-ASER-2016.pdf

Aslam, M., Rawal, S., & Saeed, S. (2017). *Public Private Partnerships (PPPs) in Education in Developing Countries: A Rigorous Review of the Evidence*. Retrieved from Education Partnerships Group website: http://arkonline.org/sites/default/files/ArkEPG_PPP_report.pdf

Atherton, P., & Kingdon, G. (2010). *The relative effectiveness and costs of contract and regular teachers in India* [CSAE Working Paper, Oxford]. Retrieved from http://www.csae.ox.ac.uk/workingpapers/pdfs/2010-15text.pdf

Avitabile, C., & de Hoyos, R. (2018). The heterogeneous effect of information on student performance: Evidence from a randomized control trial in Mexico. *Journal of Development Economics*. https://doi.org/10/gftrxv

Azam, M., Chin, A., & Prakash, N. (2013). The Returns to English-Language Skills in India. *Economic Development and Cultural Change*, *61*(2), 335–367. https://doi.org/10/gftrp9

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., … Walton, M. (2016). *Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of "Teaching at the Right Level" in India* (Working Paper No. 22746). https://doi.org/10.3386/w22746

Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L. L., & Raju, D. (2013). *Expanding Educational Opportunities in Remote Parts of the World: Evidence from a RCT of a Public-Private Partnership in Pakistan*.

Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L., Raju, D., & Ryan, S. P. (2017). *Delivering education to the underserved through a public-private partnership program in Pakistan*. National Bureau of Economic Research.

Barrera-Osorio, F., Galbert, P. de, Habyarimana, J., & Sabarwal, S. (2016). *Impact of Public-Private Partnerships on Private School Performance: Evidence from a Randomized Controlled Trial in Uganda*. Retrieved from World Bank website: http://documents.worldbank.org/curated/en/470851480966927631/Impact-of-public-private-partnerships-on-private-school-performance-evidence-from-a-randomized-controlled-trial-in-Uganda

Barrera-Osorio, F., & Raju, D. (2015). Evaluating the Impact of Public Student Subsidies on Low-Cost Private Schools in Pakistan. *The Journal of Development Studies*, *51*(7), 808–825. https://doi.org/10/gftrw8

Bau, N., & Das, J. (2017). *The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labour Market for Teachers*. Presented at the Policy Research Working Paper 8050. https://doi.org/10/gftsqm

Beasley, E. R., & Huillery, E. (2017). *Willing but unable? Short-term experimental evidence on parent empowerment and school quality* (No. WPS8125; pp. 1–78). Retrieved from The World Bank website: http://documents.worldbank.org/curated/en/982571498578771789/Willing-but-unable-short-term-experimental-evidence-on-parent-empowerment-and-school-quality

Bengali, K. (1999). *History of educational policy making and planning in Pakistan*. Sustainable Development Policy Institute.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10/ck2cfn

Bhattacharya, U. (2013). Mediating inequalities: Exploring English-medium instruction in a suburban Indian village school. *Current Issues in Language Planning*, *14*(1), 164–184. https://doi.org/10/gftrpz

Blimpo, M. P., Evans, D. K., & Lahire, N. (2015). *Parental Human Capital and Effective School Management*. Retrieved from http://www-wds.worldbank.org/servlet/WDSContentServer/IW3P/IB/2015/04/14/090224b082db7ea1/2_0/Rendered/PDF/Parental0human0ence0from0The0Gambia.pdf

Bloom, N., Genakos, C., Sadun, R., & Van Reenen, J. (2012). Management practices across firms and countries. *The Academy of Management Perspectives*, *26*(1), 12–33. https://doi.org/10/gfb7c8

Bloom, N., Lemos, R., Sadun, R., Scur, D., & Reenen, J. V. (2014). *The New Empirical Economics of Management* (CEP Occasional Paper No. 41). Retrieved from Centre for Economic Performance, LSE website: https://ideas.repec.org/p/cep/cepops/41.html

Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does Management Matter in schools? *The Economic Journal*, *125*(584), 647–674. https://doi.org/10/gbf64w

Bloom, N., Propper, C., Seiler, S., & Van Reenen, J. (2016). The Impact of Competition on Management Quality: Evidence from Public Hospitals. *Review of Economic Studies*. Retrieved from http://cep.lse.ac.uk/textonly/_new/staff/vanreenen/pdf/2016_restud.pdf

Bloom, N., Sadun, R., & Reenen, J. V. (2012). Americans Do IT Better: US Multinationals and the Productivity Miracle. *American Economic Review*, *102*(1), 167–201. https://doi.org/10/gftrv4

Bloom, N., Sadun, R., & Van Reenen, J. (2016). *Management as a Technology?* [Harvard Business School Strategy Unit Working Paper]. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2788794

Bold, T., Filmer, D., Martin, G., Molina, E., Stacy, B., Rockmore, C., … Wane, W. (2017). Enrollment without Learning: Teacher Effort, Knowledge, and Skill in Primary Schools in Africa. *Journal of Economic Perspectives*, *31*(4), 185–204. https://doi.org/10/gcptw4

Bonilla-Angel, J. D. (2011). *Contracting out Public Schools and Academic Performance: Evidence from Colombia* (Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy, University of Maryland). Retrieved from http://drum.lib.umd.edu/handle/1903/11989

Buhl-Wiggers, J., Kerwin, J., Smith, J., & Thornton, R. (2016). *The impact of teacher effectiveness on student learning in Africa*. Presented at the CSAE Conference 2017.

Carneiro, P. M., Das, J., & Reis, H. (2016). *The Value of Private Schools: Evidence from Pakistan* (SSRN Scholarly Paper No. ID 2786044). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2786044

Chakraborty, T., & Bakshi, S. K. (2016). English language premium: Evidence from a policy experiment in India. *Economics of Education Review*, *50*(C), 1–16. https://doi.org/10/f8gjbn

Chandra, A., Finkelstein, A., Sacarny, A., & Syverson, C. (2016). Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector. *American Economic Review*, *106*(8), 2110–2144. https://doi.org/10/f8xxfh

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. Retrieved from National Center for Educational Statistics website: https://files.eric.ed.gov/fulltext/ED012275.pdf

Damera, V. K. (2017). *Value addition at schools- Experimental evidence on private school and high-fee school impacts* [DPhil Thesis Blavatnik School of Government, University of Oxford].

Das, J., Dercon, S., Habyarimana, J., & Krishnan, P. (2007). Teacher Shocks and Student Learning: Evidence from Zambia. *The Journal of Human Resources*, *42*(4), 820–862. https://doi.org/10/gftrvb

Das, J., & Zajonc, T. (2010). India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, *92*(2), 175–187. https://doi.org/10/dfvbjv

Davies, E. (2017). *The Lessons Private Schools Teach: Using a Field Experiment to Understand the Effects of Private Services on Political Behavior*. Retrieved from https://sites.sas.upenn.edu/sites/default/files/emda/files/davies_emmerich_-_downstream_experiment.pdf

de Hoyos, R., Garcia-Moreno, V. A., & Patrinos, H. A. (2015). *The Impact of an Accountability Intervention with Diagnostic Feedback* [World Bank Working Paper]. Retrieved from https://www.openknowledge.worldbank.org/handle/10986/22436

Deming, D. J. (2014). Using School Choice Lotteries to Test Measures of School Effectiveness. *American Economic Review*, *104*(5), 406–411. https://doi.org/10/gftrv2

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640. https://doi.org/10/gccz25

Desai, S., Dubey, A., Vanneman, R., & Banerji, R. (2008). Private Schooling in India: A New Educational Landscape. *India Policy Forum*, *5*(1), 1–58.

Dobbie, W., & Fryer, R. G. (2013). Getting Beneath the Veil of Effective Schools: Evidence From New York City. *American Economic Journal: Applied Economics*, *5*(4), 28–60. https://doi.org/10/gftrts

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, *101*(5), 1739–1774. https://doi.org/10/bg9rzj

Duflo, E., Dupas, P., & Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of Public Economics*, *123*, 92–110. https://doi.org/10/f68tj3

Economic Policy Research Centre (EPRC). (2016). *Evaluation of the PEAS Network under the Uganda Universal Secondary Education (USE) Programme*.

Endow, T. (2018). Inferior Outcomes: Learning in Low-cost English-medium Private Schools—A Survey in Delhi and National Capital Region. *Indian Journal of Human Development*, *12*(1), 1–19. https://doi.org/10/gftrp2

Epple, D., Romano, R., & Zimmer, R. (2016). Charter Schools: A Survey of Research on Their Characteristics and Effectiveness. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 5, pp. 139–208). https://doi.org/10.1016/B978-0-444-63459-7.00003-8

Epple, Dennis, Romano, R. E., & Urquiola, M. (2017). School Vouchers: A Survey of the Economics Literature. *Journal of Economic Literature*, *55*(2), 441–492. https://doi.org/10/gftrzt

Eyles, A., Hupkau, C., & Machin, S. (2016). Academies, charter and free schools: Do new school types deliver better outcomes? *Economic Policy*, *31*(87), 453–501. https://doi.org/10/gftrt2

Eyles, A., Machin, S., & McNally, S. (2017). Unexpected school reform: Academisation of primary schools in England. *Journal of Public Economics*. https://doi.org/10/gcnn7s

Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure Data—Or tears: An application to educational enrollments in states of india*. *Demography*, *38*(1), 115–132.

Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of Technology Adoption. *Annual Review of Economics*, *2*(1), 395–424. https://doi.org/10/fh2tqm

Fryer, R. G. (2017). *Management and Student Achievement: Evidence from a Randomized Field Experiment*. Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w23437

Garlick, R., Orkin, K., & Quinn, S. (2016). *Call Me Maybe: Experimental Evidence on Using Mobile Phones to Survey African Microenterprises*. Retrieved from https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=CSAE2016&paper_id=995

Glewwe, P., & Muralidharan, K. (2015). *Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications* [RISE Working Paper]. Retrieved from http://www.riseprogramme.org/sites/www.rise.ox.ac.uk/files/RISE_WP-001_Glewwe_Muralidharan.pdf

Glewwe, P. W., Hanushek, E. A., Humpage, S. D., & Ravina, R. (2011). *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010*. Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w17554

Government of Pakistan Ministry of Federal Education and Professional Training. (2017). *National Education Policy*.

Guarino, C., Reckase, M., Stacy, B., & Wooldridge, J. (2015). A Comparison of Student Growth Percentile and Value-Added Models of Teacher Performance. *Statistics and Public Policy*, *2*(1), e1034820. https://doi.org/10/gftrs9

Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, *37*(C), 204–212. https://doi.org/10/f5n2h4

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, *104*, 212–232. https://doi.org/10/gftrsn

Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327. https://doi.org/10/chsbrh

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review*, *100*(2), 267–271. https://doi.org/10/cdb5gk

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, *73*(C), 103–130. https://doi.org/10/f6z4k8

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics*. https://doi.org/10/b37jm2

Hussain, I. (2015). Subjective Performance Evaluation in the Public Sector. *Journal of Human Resources*, *50*(1), 189–221.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10/ccfxnq

Jack, K. (2011). *Constraints on the adoption of agricultural technologies in developing countries*. Retrieved from http://sites.tufts.edu/kjack/files/2011/08/ATAI-white-paper-12102011.pdf

Jacob, B., McCall, B., & Stange, K. (2017). College as Country Club: Do Colleges Cater to Students' Preferences for Consumption? *Journal of Labor Economics*, *36*(2), 309–348. https://doi.org/10/gftrxx

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. Research Paper. MET Project. *Bill & Melinda Gates Foundation*. Retrieved from http://eric.ed.gov/?id=ED540959

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation*. Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w14607

Kim, J., Alderman, H., & Orazem, P. F. (1999). Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program. *World Bank Economic Review*. https://doi.org/10/gftrz5

Kingdon, G. G. (2007). The progress of school education in India. *Oxford Review of Economic Policy*, *23*(2), 168–195. https://doi.org/10/dc4ks7

Lassibille, G. (2016). Improving the management style of school principals: Results from a randomized trial. *Education Economics*, *0*(0), 1–21. https://doi.org/10/gftrww

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, *76*(3), 1071–1102. https://doi.org/10/d4ksrn

Lemos, R., & Scur, D. (2016). *Developing Management: An expanded evaluation tool for developing countries*.

MacLeod, W. B., & Urquiola, M. (2015). Reputation and School Competition. *American Economic Review*, *105*(11), 3471–3488. https://doi.org/10/f7wptx

Malik, R., Bari, F., Muzaffar, I., & Khan, M. S. (2015). *Partnerships for Management: Evidence from Punjab and Sindh*. Retrieved from Institute of Development and Economic Alternatives (IDEAS) website:

https://www.academia.edu/15328083/Partnerships_for_Management_Evidence_from_Punjab_and_Sindh

McKinsey. (2017, March 30). *Performance of PSSP schools—An independent perspective*. Punjab Education Sector Programme II.

Mizala, A., & Urquiola, M. (2013). School markets: The impact of information approximating schools' effectiveness. *Journal of Development Economics*, *103*, 313–335. https://doi.org/10/gftrwx

Moore, R. (2017). *Private schools in India: More learning, more inequality?* [Paper presented at 'School Effectiveness in Ethiopia, India, Vietnam and Peru: Launch of Key Findings from the 2016-17 Young Lives School Surveys'''']. University of Oxford.

Morris, C. N. (1983). Parametric Empirical Bayes Inference: Theory and Applications. *Journal of the American Statistical Association*, *78*(381), 47–55. https://doi.org/10/cdxczm

MOSPI - NSSO. (2015). *Key Indicators of Social Consumption in India—NSS 71st Round*.

Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, *31*(2), 87–106. https://doi.org/10/gdg2v9

Muralidharan, K., & Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, *119*(1), 39–77. https://doi.org/10/fgw3t7

Muralidharan, K., & Sundararaman, V. (2015). The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India. *The Quarterly Journal of Economics*, *130*(3), 1011–1066. https://doi.org/10/f7nvd2

Nguyen, Q., & Raju, D. (2014). *Private school participation in Pakistan* [World Bank Policy Research Working Paper].

Oster, E. (2016). Unobservable Selection and Coefficient Stability: Theory and Validation. *Journal of Business & Economic Statistics*. https://doi.org/10/gdz4td

Pritchett, L. (2013). *The rebirth of education: Schooling ain't learning*. Washington D.C: Center for Global Development.

Pritchett, L., & Aiyar, Y. (2014). *Value Subtraction in Public Sector Production: Accounting versus Economic Cost of Primary Schooling in India* (SSRN Scholarly Paper No. ID 2623073). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2623073

Pritchett, L., & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skill levels. *International Journal of Educational Development*. https://doi.org/10/f65t5b

PROBE Team. (1999). *Public report on basic education in India*. Retrieved from Oxford University, New Delhi, IN. website: http://www.undp.org/content/dam/india/docs/public_report_basic_education_india.pdf

Rao, G. (2018). Familiarity does not breed contempt: Diversity, discrimination and generosity in Delhi schools. *American Economic Review*. Retrieved from http://economics.cornell.edu/sites/default/files/files/events/Gautam.Rao_JMP_latest.pdf

Romero, M., Sandefur, J., & Sandholtz, W. A. (2017). Can Outsourcing Improve Liberia's Schools? Preliminary Results from Year One of a Three-Year Randomized Evaluation of Partnership Schools for Liberia. *Working Paper 462*. https://doi.org/10/gftr4n

Sandefur, J. (2013). Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education - Working Paper 321. *Center For Global Development*. Retrieved from http://www.cgdev.org/publication/scaling-what-works-experimental-evidence-external-validity-kenyan-education-working

Sarin, A., Kuhn, S., Singh, B. D., Khanghta, P., Dongre, A. A., Joshi, E., … Rahman, F. (2015). *State of the Nation: RTE Section 12 (1)(c)*. https://doi.org/10/gftrwr

Shakeel, M., Anderson, K. P., & Wolf, P. J. (2016). *The participant effects of private school vouchers across the globe: A meta-analytic and systematic review* [University of Arkansas Department of Education Reform Working Paper]. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2777633

Singh, A. (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, *113*, 16–32. https://doi.org/10/gftrsm

Srivastava, P. (2006). Private schooling and mental models about girls' schooling in India. *Compare: A Journal of Comparative and International Education*, *36*(4), 497–514. https://doi.org/10/d92jmq

Tavares, P. A. (2015). The impact of school management practices on educational performance: Evidence from public schools in São Paulo. *Economics of Education Review*, *48*, 1–15. https://doi.org/10/f73zpc

Termes, A., Bonal, X., Verger, A., Zancajo, A., López, L., Ramírez, Y. C., & Sierra, A. (2015). *Public-Private Partnerships in Colombian Education: The Equity and Quality Impications of "Colegios en concesión"*. Retrieved from https://pure.uva.nl/ws/files/2599921/175603_WorkingPaper66_PPPs.pdf

The Economist. (2015, August 1). Learning unleashed: Where governments are failing to provide youngsters with a decent education, the private sector is stepping in.

Technical report. *The Economist*. Retrieved from https://www.economist.com/briefing/2015/08/01/learning-unleashed

The World Bank. (2013). *World Development Indicators*. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement*. *The Economic Journal*, *113*(485), F3–F33. https://doi.org/10/bgzbmj

Tooley, J. (2016). Extending access to low-cost private schools through vouchers: An alternative interpretation of a two-stage 'School Choice' experiment in India. *Oxford Review of Education*, *42*(5), 579–593. https://doi.org/10/gftrz2

Tooley, J., & Dixon, P. (2007). Private schooling for low-income families: A census and comparative survey in East Delhi, India. *International Journal of Educational Development*, *27*(2), 205–219. https://doi.org/10/dvpp6m

Uganda Bureau of Statistics (UBOS). (2014). *Uganda National Household Survey 2012/2013*. Kampala Uganda: UBOS.

Ugandan Ministry of Education and Sport, & UNESCO - IIEP Pôle de Dakar. (2014). *Teacher Issues in Uganda: A shared vision for an effective teachers policy*. Retrieved from http://unesdoc.unesco.org/images/0022/002297/229777e.pdf

UNESCO. (2014). *Teaching And Learning: Achieving quality for all*. Retrieved from UNESCO Publishing website: http://unesdoc.unesco.org/images/0022/002266/226662e.pdf

West, M. R., Buckley, K., Krachman, S. B., & Bookman, N. (2017). Development and implementation of student social-emotional surveys in the CORE Districts. *Journal of Applied Developmental Psychology*. https://doi.org/10/gdn2dc

Westfall, P. H., Young, S. S., & Wright, S. P. (1993). On Adjusting P-Values for Multiplicity. *Biometrics*, *49*(3), 941. https://doi.org/10/brfgvd

Wolf, P. J., Egalite, A. J., & Dixon, P. (2015). Private school choice in developing countries: Experimental results from Delhi, India. In *Handbook of International Development and Education* (p. 456). Retrieved from http://books.google.com/books?hl=en&lr=&id=_-rECQAAQBAJ&oi=fnd&pg=PA456&dq=%22to+enroll+their+child+in+a+private+school+of+their+choosing+so%22+%22topics+in+education+policy+are+more+contentious+than%22+%22better-%C2%ADmanaged+than+government+schools%3B+(2)+by+providing%22+&ots=r25GIxeORQ&sig=f5JlsThgPZYVSOL0HdXM2Rz9wjg

World Bank. (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Washington D.C.: World Bank Publications.

Young, A. (2019). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *The Quarterly Journal of Economics*, *134*(2), 557–598. https://doi.org/10.1093/qje/qjy029