Learning in Lagos: Comparing Student Achievement in Bridge, Public, and Private Schools.

EDOREN
Education Data, Research and Evaluation in Nigeria

UKaid
from the British people

Acknowledgements

This report owes its deepest debt to pupils, parents, teachers, and head teachers for their time and patience in participating in this research.

Femi Adegoke (Oxford Policy Management Nigeria Country Director), Gloria Olisenekwu (Oxford Policy Management Nigeria), Bridget Azubuike (independent consultant), alongside forty five talented enumerators have played a key role in the implementation of the survey.

This report was written by Alina Lipcan (Oxford Policy Management Education Portfolio), Lee Crawfurd (University of Sussex), and Brian Law (independent consultant). Ian MacAuslan (Oxford Policy Management Education Portfolio Leader) and Michele Binci (Oxford Policy Management Senior Impact Evaluation Consultant) provided quality assurance to the study's design and analysis.

Many thanks are due to Ed Smithson (UK Department for International Development), Chris Okeke (UK Department for International Development), Esohe Eigibike (UK Department for International Development), Monazza Aslam (University College London Institute of Education), Justin Sandefur (Centre for Global Development), Lisa Chen (Bridge) and Steve Cantrell (Bridge) for helpful comments on drafts of the report.

All errors remain the responsibility of the authors.

We are grateful to UKAid for funding this study.

Executive summary

Despite the relatively small number of public schools in Lagos, only 4% of primary school aged children were out of school in 2011 (World Bank, 2017). High enrolment rates are enabled by private schools filling the gap in provision. In 2011, 57% of primary and secondary students in Lagos studied in 12,098 private schools (Härmä, 2013). Many of these schools can be considered to be "low-fee" – affordable even to those on the poverty line. This comparative study is designed to describe the learning levels of students in Bridge, public, and private schools in Lagos and identify factors that may help account for differences in achievement. The study is best considered as an examination of the current state of these schools in Lagos and a baseline for future work. The study is the first step in assisting the Department for International Development (DFID) to understand the value of its investment in supporting Bridge to enter the Lagos market. Bridge received a grant through DFID's Developing Effective Private Education Nigeria (DEEPEN) Innovation Fund to enter Lagos. The DEEPEN Innovation Fund aimed to increase competition in the low-fee education market and to improve quality in low-fee schools in Lagos.

Research Questions

1. How does the literacy and numeracy performance of Primary 2 (P2) students compare between Bridge, public, and private schools?

2. How do teacher and school characteristics (e.g. teaching practice, school management, and school fees) vary within and among school types?

3. To what extent are pupil, teacher, and school- level factors correlated with student achievement?

This study is not designed to determine causality – the effect of a particular school type on learning outcomes**.** Neither are we able to make a rigorous statement about the value for money of DFID's investment, as this would require both a causal estimate of the effectiveness of Bridge schools, and an estimate of the eventual number of Bridge schools in Lagos. We do provide a comparison of test scores and fees charged to parents across school types.

Results

1. Literacy, numeracy, and value for money

- In literacy, students at Bridge schools have better performance than students at other private schools (by 0.35 standard deviations) and public schools (by 1.38 standard deviations);

- In mathematics, students at Bridge schools have similar performance to students at other private schools and better performance than students at public schools (by 0.86 standard deviations);

- Adjustments for pupil age and gender, household wealth and home language, and local government area reduce differences, particularly in private schools - by 0.12 in literacy compared to private schools, by 0.01 in literacy compared to public schools, and by 0.03 in numeracy compared to public schools. However, they remain sizeable.

  Our findings suggest that Bridge may provide parents with better value for money in literacy. This does not necessarily equate with better value for money from a policymakers' perspective – determining whether investments provide taxpayers with value for money requires an impact estimate and more information on the cost and financial modelling of both public and private provision.

2. Students, teaching and management

- Students in Bridge schools and other private schools are less likely than students in public schools to belong to the bottom socioeconomic quintiles, to speak Yoruba at home, and to be overage.
- All school types enrol equal proportions of boys and girls.
- Teachers in Bridge schools report higher motivation than teachers in other schools, and Bridge schools are better managed than other schools. However, observed teaching practice does not differ substantially between school types for literacy. For numeracy, teachers in Bridge and public schools spend more time teaching than teachers in private schools.

3. Correlations with learning

- For literacy, students from better socioeconomic backgrounds have higher learning achievement in private and public schools, but not at Bridge schools. Numeracy achievement is not significantly correlated with wealth across school types.
- Girls perform significantly better than boys overall, and particularly in Bridge and public schools. The difference is fairly large – 0.16 standard deviations overall. In private schools, girls and boys perform similarly at the P2 level, in line with previous evidence from Lagos (EDOREN, 2016). We also find a statistically significant disadvantage for boys in literacy in public schools at the P4 level (EDOREN, 2017).
- Better school management is significantly correlated with higher literacy and numeracy achievement, particularly in Bridge schools and public schools.
- Teaching qualifications and bachelors' degrees are not significantly correlated with literacy and numeracy achievement.

Recommendations

- This study is not designed to estimate the causal effect of attending a Bridge school. A follow-up survey tracking the same pupils over time would allow us to estimate pupil progress and make a claim to causal inference.
- Further research should seek to understand why the stark differences in numeracy achievement in private schools compared to public schools found in this report are not as pronounced at the Primary 4 (P4) level (EDOREN, 2017).
- Further research should also seek to understand the reasons why girls in Lagos have higher literacy rates compared to boys.

Table of contents

List of figures and tables

List of abbreviations

DFID	Department for International Development
DEEPEN	Developing Effective Private Education in Nigeria
EDOREN	Education Data, Research and Evaluation in Nigeria
ESSPIN	Education Sector Support Programme in Nigeria
IRT	Item response theory
LGA	Local Government Area
NGN	Nigerian Naira
OLS	Ordinary least squares
PCA	Principal component analysis
P2	Primary Grade 2
P4	Primary Grade 4

1 Introduction

With 1,200 primary-age children for every available public primary school, Lagos state has the fewest number of public schools per capita in Nigeria (World Bank, 2017). Lagos is one of the fastest growing megacities in the world, with its population due to increase two-fold by 2050 according to some predictions (World Bank, 2017). Thus, pressure on meeting the need of educating the next generation can only be expected to intensify. Despite the relatively low number of public schools, only 4% of primary school aged children were out of school in 2011 (World Bank, 2017). High enrolment rates are enabled by private schools filling the gap in provision. In 2011, 57% of primary and secondary students in Lagos studied in 12,098 private schools (Härmä, 2013). Private schools in Lagos do not cater exclusively for the wealthy: recent research suggests that 69% of those attending low-cost private schools and 56% of those attending medium-cost in Lagos can be classified as poor – i.e. living below the 2013 Lagos poverty line of NGN 112,895 Naira per annum ($313 at current exchange rates) (Yngstrom, 2014; Tooley J. , 2013).

DFID Nigeria recognises that the private sector, as well as the public sector, contributes to the education of Lagosian children and that DFID therefore needs to work with the private sector as well as the public sector to improve education outcomes for children in Lagos (DFID 2011). DFID funds the DEEPEN programme that seeks to make education markets in Lagos work for the poor. As part of the DEEPEN programme, there was an innovative business model fund and a procurement process, that sought to increase competition and quality service provision of low-fee schools in Lagos. Bridge International Academies (Bridge) received an outcome-based contract in 2014 to enter the Lagos market as part of the DEEPEN programme Innovation Fund.[1]

This study is the first step to help DFID understand the quality of education provided by Bridge, relative to alternatives. The study aims to answer the following questions, in each case comparing Bridge schools with other private and public schools:

1. What are the levels of learning for P2[2] students at these schools?
2. How do student, teacher and school characteristics (e.g. student background, teaching practice, school management, and school fees) vary within and among school types?
3. To what extent are pupil, teacher, and school- level factors correlated with student achievement?

This comparative study is designed to describe learning levels for public, private, and Bridge schools in Lagos and identify factors that may help account for differences in learning achievement[3]. This study is best considered as an examination of the current state of these schools in Lagos and a baseline for future work. It is not designed to determine causality – the effect of a particular school type on learning outcomes. With one observational survey and non-experimental assignment of students to schools, any attempt to attribute identified differences in student learning levels to schools will conflate the influence of schools and teachers with the influence of student background.

An additional survey round tracking the same students over time would greatly improve our ability to control for (time-invariant) unobserved pupil and family background characteristics, thus allowing us to be make a stronger claim to causal estimates of the effect of attending different school types, something we do not claim to be able to do here. A second survey round would take place in May/June 2020 when panel students are at the end of P4, and allow us to estimate pupil growth (value-added) models controlling for both observable and unobservable differences between pupil backgrounds, which more clearly isolates the

[1] Separately, as the capital source for the Commonwealth Development Corporation (CDC) and as a Limited Partner in NovaStar, DFID has invested in Bridge's global entity
[2] This level of education was chosen to maximise the sample size of students available, as not all Bridge schools currently offer education at P3, P4, and P5
[3] Throughout this report, we refer to non-Bridge low and medium fee schools as private.

effect of schools. The extension will be necessary for this study to help resolve important questions such as "to what extent are observed differences between public and private school student achievement *due* to student background characteristics?" (Ashley, 2014).

2 Approach and methods

2.1 Conceptual framework

We adopt a simple theoretical framework where learning outcomes are influenced by characteristics of pupils, their households, their teachers, and their schools. Our data collection and analysis is designed to capture as much as possible these relevant factors.

2.1.1 Pupil and household level factors

A standard predictor of learning outcomes are students' individual characteristics, such as age, gender (observable) and levels of academic motivation (unobservable). These are separate from, but related to students' family background (e.g. language spoken at home and socioeconomic status and parental education), which also influence learning achievement. Research in Lagos has demonstrated that children from wealthier households have higher levels of literacy and numeracy at the P2 level (EDOREN, 2016), and the Primary 4 (P4) level (EDOREN, 2017). Speaking English at home is correlated with better literacy and numeracy scores (EDOREN, 2016). Previous evidence has also shown that, accounting for pupil background characteristics, girls perform better than boys in literacy at the P4 level in both private and public schools in Lagos (EDOREN, 2017). Underage students have also been found to be significantly more likely to have lower numeracy achievement (EDOREN, 2016).

Given differences in student background characteristics have been associated with different levels of learning, we have collected data on socioeconomic background, language spoken at home, gender, and age. However, we will not be able to directly account for unobservable characteristics (child, family or school-level) that influence learning and because some of these characteristics are likely to influence the choice of schooling, the resultant analysis is unable to fully establish whether differences in learning outcomes are due to a child being in a given school type.

2.1.2 Teacher level factors

Teaching quality is the most important institutional influence on student outcomes. Several studies show the need for interventions that focus on teachers and teaching quality (Burgess, Davies, & Slater, 2009; Hanushek & Woessmann, 2011; Singh & Sarkar, 2012). Drivers of teacher effectiveness are varied – including teacher qualifications and training, teacher competence and subject-knowledge, as well as effective management (Aslam, et al., 2016) and teacher practice whilst in class (Aslam and Kingdon 2011). Previous research in Lagos has found that motivation in combination with high qualifications (university degree or higher) are correlated with better student learning outcomes in literacy, but not in numeracy (EDOREN, 2016). In the absence of high levels of motivation, having highly qualified teachers is not correlated with learning.

In this study we explore the influence of teaching practices through lesson observations gathering data on teacher time on task and specific teaching practices, along with questions on teacher motivation.

2.1.3 School level factors

Characteristics of schools that may influence children's learning include the quality of school infrastructure, the availability of learning materials and resources, and school leadership and management. Several recent literature reviews (e.g. Muralidharan, 2013) underscore the limited impact that improvements in school infrastructure, toilets, electricity and other educational inputs have had on learning achievements. Previous work in Lagos finds some correlation between top quality infrastructure and literacy, but no significant

correlation with numeracy (EDOREN, 2016). Given this, we collect limited data on school infrastructure in this study.

We also measure the fees charged by schools and the additional costs charged to parents. Previous research in Lagos suggests that once the effect of socioeconomic background is accounted for, higher fees are not significantly correlated with learning in low fee schools (which are also the subject of this research) (EDOREN, 2017). However, given the scarcity of evidence on the relationship between fees and learning, and the importance of the relationship for both parents' and policymakers' decision-making, we collect detailed data on fees from parents and head teachers. We do not propose to estimate the full economic cost of provision of different types of schools (including for example government and philanthropic spending), as this is hampered both by the availability of historic and/or commercially sensitive data, and conceptual difficulties in the allocation of fixed central administrative and upfront investment costs to individual students and schools.

Finally, we pay close attention to management practices at the school level. Recent studies in Uganda (Crawfurd, 2017), India (Lemos & Scur, 2017), and Liberia (Romero, Sandefur, & Sandholtz, 2017) uncover strong and significant relationships between school leadership and management and learning outcomes. Therefore we include an adapted version of the World Management Survey tool from Crawfurd (2017) in our research. We also ask about the frequency of external and internal lesson observation and pay close attention to head teachers' qualifications and levels of management experience.

2.2 Instruments

Table 1 Instruments used

Concepts	Instruments
Student Learning	Literacy & Numeracy Assessment
Student Characteristics	Student survey
Household Characteristics	Household Wealth Index
Teacher Characteristics	Classroom Observation & Teacher motivation questionnaire
School Characteristics	School fees & School Management questionnaire

2.2.1 Students

Literacy and numeracy measurement

To measure learning achievement, numeracy and literacy test instruments were adapted in order to ensure comparisons could be made across the various EDOREN learning assessments and evaluations in Nigeria. To avoid the ceiling effects observed in the DEEPEN baseline survey, five new items testing higher order literacy skills such as writing and comprehension were added to the literacy test. Five new items were also added to the numeracy test to insure against ceiling effects. The new tests were piloted in November 2017 on a sample including 90 children from Bridge, low and medium fee private schools, and public schools. For both the pilot and the full survey, the test questions were programmed on tablets and pupil books were used for rough calculations or to write down answers. OPM staff stored the tablets at the end of each day of data collection. All the pupil books from EDOREN assessments, including this one, are collected and safely stored in OPM's Abuja office. OPM works with a set of trusted and professional enumerators who are briefed to never share tests questions or test booklets.

The pilot data was analysed using item response theory. Item Response Theory (IRT) improves on raw test scores (i.e. percentage of correct answers) by contextualising test results to reflect both children's

proficiency levels and the difficulty of individual test items. Proficiency and item difficulty are estimated through an iterative process of model-fitting and allows items and individuals taking the test to be placed on a scale reflecting whether test-takers were able to provide correct answers to easy, average, and difficult test items. As noted by Das and Zajonc (2010), IRT is routinely used in education research and in most large-scale testing situations.

The pilot tests showed no ceiling effects (i.e. a high percentage of students obtaining the top score) or floor effects (i.e. a high percentage of students obtaining a zero score). Most of the items on the test distinguished levels of student performance, and the pilot data separated means for Bridge, private and public schools. The full survey took place between the 15th of January 2018 and the 4th February 2018, at the beginning of the second term of their Primary 2 year.

The IRT analysis finds that for both literacy and numeracy there are no obvious floor effects. There are no ceiling effects for numeracy, but there is a small ceiling effect for literacy (i.e. 6% of children in the sample obtained the top score: 0% in public schools, 5% in private schools, and 12% at Bridge). Fortunately, as discussed in the rest of the report, the small ceiling effect does not affect in any substantive way our ability to draw conclusions from the data. This is because differences in literacy performance between the three types of schools are pronounced and statistically significant even when controlling for a battery of background characteristics. The literacy premium between Bridge schools and other schools types, and that between private schools and public schools is likely to be underestimated due to the ceiling effect. In a longitudinal study at the P4 level, we will be able to further adapt the test to take into account ceiling effects using an existing instrument developed for the P4 level. In addition, data modelling techniques such as the Bayesian Tobit growth curve model will be employed to account for ceiling effects (Wang, Zhang, McArdle, & Salthouse, 2008). Annex A contains more details on our approach to measuring learning.

Household wealth index

Household wealth is typically both strongly correlated with student performance, and may vary systematically between types of school. By measuring household wealth, we can present between school type comparisons of learning levels after adjusting for household wealth. This data also helps to understand the overall level of welfare of households whose children attend private and public schools.

A detailed survey of household food and non-food expenditures was not possible as the complexity of these types of expenditure rules out the possibility of administering such a survey to children. The alternative option of administering questions on assets to children's parents carried the risk of enumerators not being able to track and survey a high proportion of the parents. Therefore the team opted to obtain a measure of socioeconomic background from a relatively simple set of questions that a child can answer and that is adapted to the Lagos-specific context to ensure that it captures enough variability in its measure. This method has been successfully validated for P2 students and used in three other EDOREN studies (DEEPEN Baseline, ESSPIN, P4 Comparison study). A technical write-up demonstrating a high level of consistency between children's and parents' responses from the DEEPEN study is available (EDOREN, 2015).

A household wealth index was derived based on coefficients derived from a polychoric principal component analysis (PCA) on our data. We use polychoric PCA instead of the simple PCA approach popularised by Filmer and Pritchett (2001) in order to correct for the fact that simple PCA with a collection of ordinal variables might produce spurious correlations (Kolenikov & Angeles, 2009). However, the essence of the technique is the same: PCA looks for the linear combinations of variables that capture the information they jointly convey most successfully. More detail on the calculation of the index is available in Annex C.

2.2.2 Teachers and Teaching

Classroom observation

The research team used the lesson observation instrument used extensively by the World Bank in sub-Saharan Africa in their Service Delivery Indicator surveys. These surveys are a simplification of the Stallings tool and measure the quantity of teaching (measuring quality requires a higher level of expertise than the typical survey enumerator available to us at short notice). This instrument provides an estimate of what percentage of lesson time is spent on task, along with some simple data about the classroom and the teacher.

Teacher motivation

Teacher motivation is measured using 29 questions on a four-level Likert scale on perceptions about various aspects of teaching. The questions have been previously piloted and refined in Nigeria as part of TDP and DEEPEN evaluations. The questions cover the teaching environment, career aspirations, teaching abilities and their day-to-day enthusiasm for teaching. Principal component analysis was used to analyse the correlations between these questions to calculate a score ranging from -5 to 5 with a mean close to 0.

2.2.3 Schools

School fees

We collected detailed information from head teachers and from parents of children attending all the school types on all expenditures from parents, whether spent at the school or outside of the school. To calculate fees, we asked head teachers[4] in private and Bridge schools and parents for mandatory fees that are paid to the school, including tuition, registration fee, exam and report fees, textbooks and writing materials, building maintenance costs, costs of uniforms, transport and additional instruction fees. Following Tooley & Yngstrom (2014), we have attempted to restrict the school fee calculations to the following: tuition fee, registration fee, exam/tests, report cards, and building development/maintenance. Where possible, we have excluded writing materials, textbooks, tissues, transport and additional instruction fees. However, enumerators recorded fees as a lump sum where parents were unable to recall the breakdown, and Bridge standard fees are inclusive of textbooks and additional instruction after school. Adjusting the fee thresholds for inflation since 2013, we classify low-cost private schools in Lagos as those costing below 42,802 Naira per year ($120), medium-cost private schools cost 42,802-85,605 ($120-$238) Naira yearly and high-cost school fees - over 85,605 Naira[5][6]. To put these fees in context, median household expenditure on private schools in Lagos adjusted for 2017 inflation was 49,770 Naira ($137) (including school fees and registration, school repairs, uniforms, books, and other expenses), according to the 2015-2016 General Household Survey. If additional tuition is included in the total, the median expenditure is 58,016 Naira per year ($160).

[4] Data about fees was not collected from head teachers in public schools. The survey team collected fees data during the preliminary visits seeking permission to visit the schools. Officially, no fees should be charged by public schools. Therefore, given the difficulties of obtaining permission to collect data in public schools the survey team decided against asking head teachers sensitive questions about fees.

[5] The inflation rates for 2013, 2014, 2015, 2016, 2017 used to make the adjustments were the All Items inflation rates obtained from official data from the Central Bank of Nigeria - https://www.cbn.gov.ng/rates/inflrates.asp?year=2014

[6] Bridge fees are inclusive of textbooks and additional instruction after school. These are included in standard fees and cannot be parsed out.

School management

We measured school management quality using a version of the Bloom et al. (2015) tool adapted for Uganda by Crawfurd (2017) and for Lagos by the research team. We carried out interviews with head teachers and administered 15 questions covering five main areas: target-setting, monitoring, operations (planning and leading teaching), people (teacher) management, and leadership. These are detailed below.

- 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 ongoing 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;
- Leadership: how the school's vision is set.

The research team adapted the tool, most notably through expanding the tool "horizontally" (by introducing half scores) to allow for greater variation of scores and finer differentiation between levels of management, following Lemos & Scur (2017) and also clarifying the scoring rubric to be more in line with the original conceptual distinctions from the World Management Survey (e.g. clarifying that a higher score should be awarded the more evidence there is of a structured system being in place). Each interview was double scored: the first interviewer was accompanied by a second interviewer whose main role was to monitoring the quality of the interview being conducted by taking notes and separately scoring the responses after the interviews had ended. Additional detail on the adaptation and implementation of the tool is available in Annex B.

2.3 Sampling

All 37 Bridge schools in Lagos were matched with low- and medium-fee private schools derived from EDOREN's recent P4 study, which compared private and public school P4 pupils in Lagos. This study contains the most comprehensive dataset for our variables of interest. Matching was conducted using a composite score index based on three variables: nominal school fees, total number of pupils, and GPS coordinates (longitude and latitude). A randomly selected sub-set of public schools closest to the sample of Bridge schools was then drawn from the P4 sample to create the public school arm for this study. The matching procedures followed are detailed in Annex D.

In each school, a P2 teacher was randomly selected from a list provided by the head teacher. This teacher's classroom was observed, and the teacher was administered the teacher motivation questionnaire. Fifteen children from the teacher's classroom were randomly assigned to sit a numeracy test or a literacy test using sheets placed in a bag. In many of the private schools sampled, teams were not able to sample and administer questionnaires to 15 pupils because schools had class sizes with fewer than 15 pupils. In order to obtain the required number of students, the team therefore visited a larger than planned number of private schools.

A total of 124 schools were visited across several LGAs in Lagos State and the following instruments were administered: Pupil test, classroom observation, teacher interview, head teacher interview, and parents' interview, as summarised in the table below.

Table 2 Sample

	Bridge	Private	Public
Schools			
Target sample schools	37	37	37
Number of schools surveyed	37	49	38
Percentage of sample schools surveyed	100%	132%	103%
Pupils			
Target sample pupils	555	555	555
Number of pupils interviewed	553	503	553
Percentage of target pupils surveyed	100%	91%	100%
Teachers			
Target sample teachers	37	37	37
Number of teachers interviewed	37	45	37
Percentage of target teachers surveyed	100%	122%	100%
Head teachers			
Target sample head teachers	37	37	37
Number of head teachers	37	49	38
Percentage of target teachers surveyed	100%	132%	103%
Classroom observation			
Target sample classroom observations	74	74	74
Number of classroom observations conducted	73	91	69
Percentage of classroom observations surveyed	99%	123%	93%
Parents			
Target sample parents	555	555	555
Number of parents interviewed	507	384	433
Percentage of target parents surveyed	91%	69%	78%

The survey team replaced five public school and 12 private schools from the original sample due to refusal to participate, fees being too high, school change of location etc. The small class sizes found in private schools, combined with the large number of refusals in private schools, led the research team to make the decision to survey all 20 eligible private schools in the replacement sample. Each private school in the replacement sample was the second-best match for Bridge schools using the index. Further details about the implementation of the survey are available in the fieldwork implementation report (Bridge Fieldwork Implementation Report, 2018).

2.4 Study governance

This study was overseen by DFID and based on transparency. A memorandum of understanding (MoU) between DFID, Bridge and EDOREN was signed, defining the research questions and study design. The MoU made explicit that Bridge have no editorial control over the study's findings. DFID, EDOREN, and Bridge had regular calls to discuss progress.

DFID's oversight of the evaluation was unchanged from baseline. The evaluation team reported to the DFID Nigeria education team on progress towards evaluation objectives through regular EDOREN quarterly written reporting to DFID, and six weekly verbal project management updates.

The governance arrangements are designed to underpin the evaluation's independence. The close relationship between EDOREN and Bridge, and the role of DFID as both guarantor of independence and funder, leaves concerns. Our approach to mitigating these concerns is to 1) develop all research outputs independently within EDOREN, building on OPM's reputation for rigour and independence. Specifically, all data collection and analysis was conducted independently by EDOREN; neither Bridge nor DFID have access to the data. 2) ensure that report drafts are peer reviewed by independent researchers (in this case Monazza Aslam). and 3) ensure data and analytical approaches are publicly available for scrutiny, with appropriate confidentiality safeguards to protect respondents.

2.5 Methodology

2.5.1 Descriptive Statistics

We begin by comparing average (mean) values across types of school, looking at student learning, household characteristics, teachers and teaching, and school management. Visual analyses and statistical tests provide an initial indication on whether differences across groups are practically meaningful and statistically significant.[7] This descriptive analysis does not control for factors that may confound the associative patterns.

2.5.2 Correlation analysis

Next we look at correlations between the different measured factors and student learning, in an Ordinary Least Squares (OLS) regression framework. Literacy and numeracy outcomes are the dependent variables, and household, teacher, and school factors are explanatory variables. The magnitude and statistical significance of the coefficients associated with these explanatory variables provide an indication of whether any descriptive association detected is still apparent when controlling for other potentially confounding pupil-, teacher-, and school-level factors.

These regression analyses are implemented for Bridge, private, and public school samples separately so as to determine the magnitude and significance levels of any detected correlation between explanatory variables and learning outcomes separately on the sub-sample of pupils studying in the different school types.

2.5.3 Coefficient stability methods

Third, we examine the size of the correlation between school type (Bridge, private, or government) and student performance, and make some initial statements about how much of this correlation may be confounded by student family background, and what remaining difference may be a causal effect of schools. We approach this using the “coefficient” stability method outlined in by Oster (2016). This method compares the correlation between school type and student performance with and without controls for confounding factors. We then make some assumptions about the likely size of unobserved confounders, to place some possible bounds on the causal effect of schools. This method is described in more detail alongside the results in Section 3.4.

[7] Tables will present estimates and standard errors by group, while graphs will include confidence intervals.

3 Results

In this section, we answer the three key research questions.

Section 3.1 describes the level of learning achievement for P2 students attending these schools in Lagos.

Section 3.2 describes how student, teacher, and school characteristics vary within and between school types.

Section 3.3 describes how student, household, teacher, and school characteristics are correlated with learning achievement. We present both simple comparisons of means across groups, and OLS multiple regressions showing the conditional correlations between student, household, teacher, and school factors with learning.

3.1 Literacy and numeracy outcomes in Bridge, private, and public schools

3.1.1 Overview of literacy and numeracy outcomes

Students at Bridge schools have higher literacy scores than students at both private and public schools. 80% of students in Bridge schools perform above the sample average (standardised as 0), compared to 62% of students in private schools and 18% in public schools (Figure 1, left panel)[8]. This difference in achievement between schools is statistically significant.

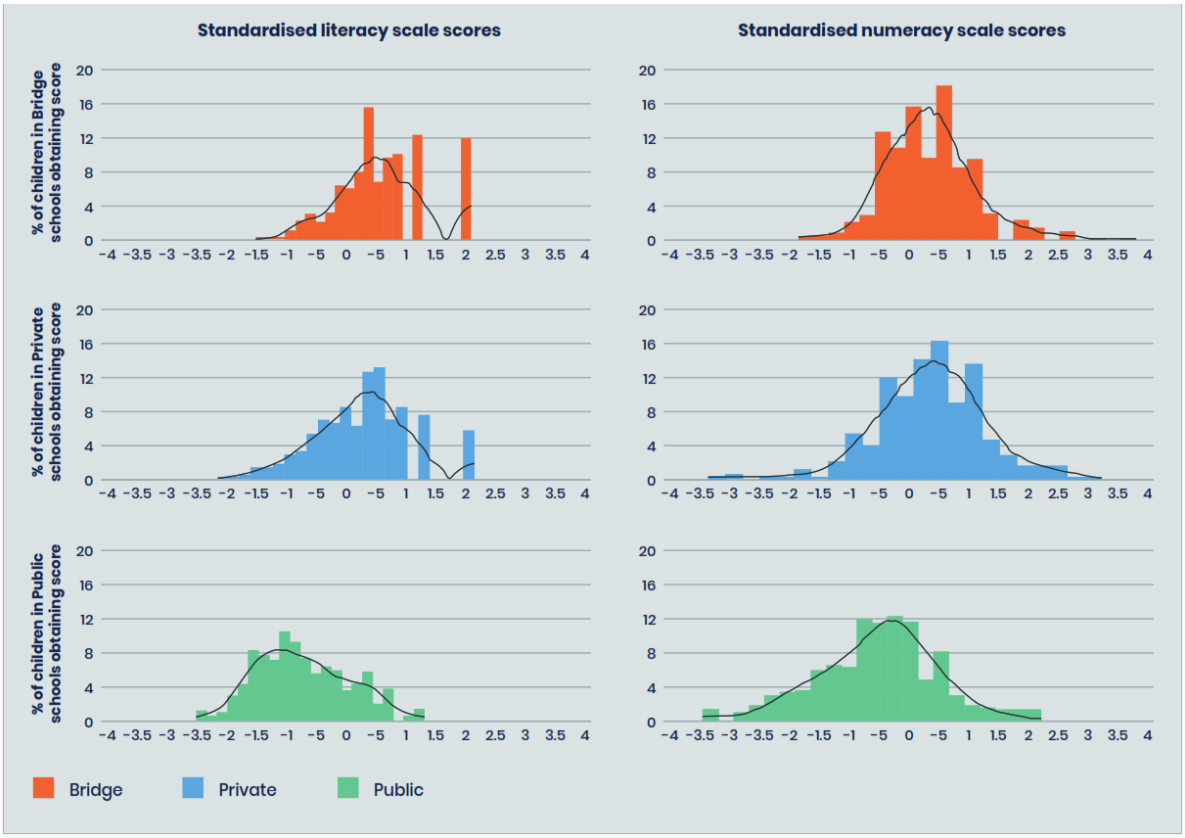
In numeracy, students at Bridge schools have higher scores than students at public schools, and similar scores to students at other private schools. 62% of students at Bridge schools perform above the sample average, compared to 64% of students at other private schools (the difference between the two school types is not statistically significant)[9]. Students at Bridge and private schools are significantly more likely to perform above average than children in public schools, where only 24% of students perform above the sample average[10].

[8] The p-values of the differences between Bridge schools and other private schools, between Bridge schools and public schools, and between private and public schools are 0.0000.

[9] The p-value of the difference between Bridge schools and other private schools is 0.2575.

[10] The p-values of the difference between Bridge schools and public schools, and between private and public schools are 0.000.

Figure 1 Distribution of literacy (left) and numeracy (right) scores in Bridge (orange), private (blue) and public (green) schools



Letter name identification was amongst the easiest literacy items in the test, and answering simple comprehension questions was amongst the most difficult. The item-level analysis below is presented for illustrative purposes only: throughout this report we use scale scores derived using item response theory to make comparisons between schools. The scale scores take into account item difficulty, not only the percentage of correct responses, when determining a pupil's level of proficiency, thus ensuring that ability rather than test quality is measured. They are derived using item response theory to make comparisons between schools.

Table 3 Literacy items: an illustration of what children can do

Literacy item examples	% answering correctly		
	Bridge	Private	Public
Identify the majority of 50 letter names and sounds in 1 minute	98	96	77
Read the majority of 25 single-syllable words in 1 minute	93	87	58
Read the majority of words in a grade 2 level written text correctly	97	86	34
Write a sentence	95	88	67
Listen to a short passage and answer simple comprehension questions	75	62	42

One-digit multiplication was amongst the easiest numeracy items in the test and the subtraction and addition of 2 or 3 digit numbers, or operations requiring carry over, were amongst the most difficult. The figure below presents item-level analysis for numeracy outcomes, which should be used to inform rather than replace conclusions based on scale scores presented in the following sections.

Table 4 Numeracy items: an illustration of what children can do

Numeracy item examples	% answering correctly		
	Bridge	Private	Public
Find the solution to 2 simple multiplication problems	98	94	84
Find the solution to at least 3 of 4 simple addition problems	87	83	61
Find the solution to at least 3 of 4 simple subtraction problems	91	88	60
Colour in three quarters of a shape and one third of a shape	55	58	43
Find the solution to at least 3 of 4 addition and subtraction problems involving two-digit numbers/three-digit numbers/carry over	44	45	14

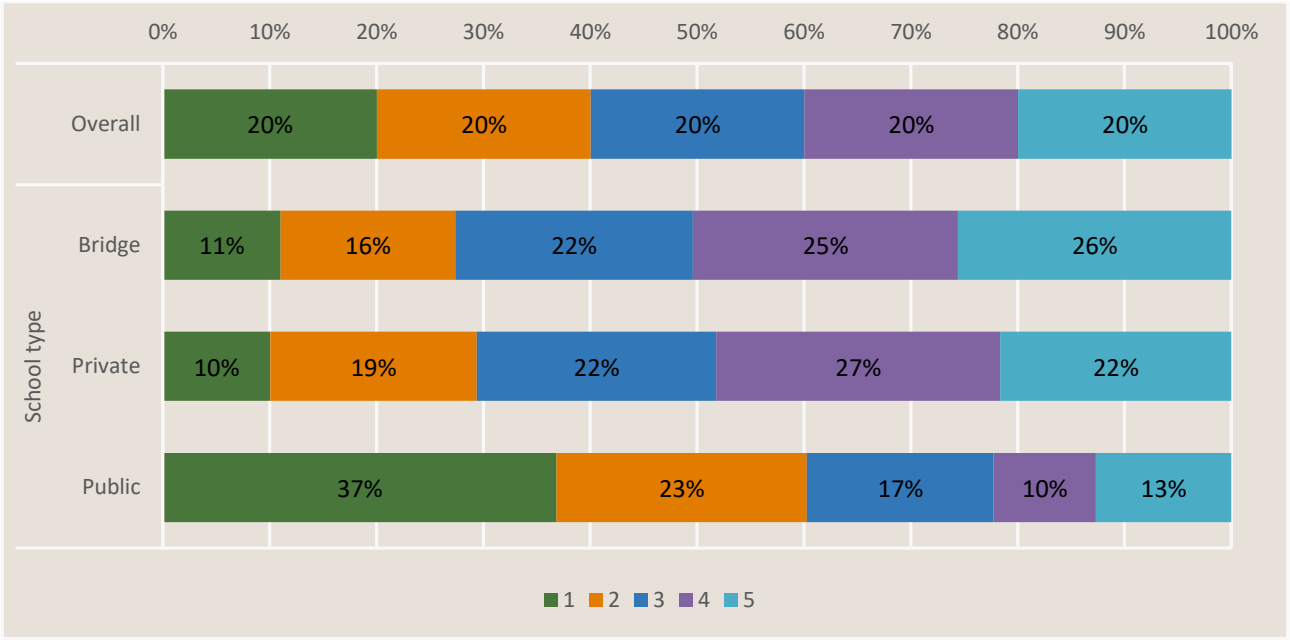
3.2 Student characteristics, teaching, school management, and school fees by school type

3.2.1 Student characteristics

Pupils in public schools tend to be somewhat poorer than students in Bridge and private schools. We assigned all of the pupils in our sample to five equally sized groups (quintiles) based on a household wealth index (comprising facilities at their house and a list of common assets). 60% of the students in public schools belong to the two bottom quintiles, compared with 27% in Bridge schools and 29% in other private schools[11]. The distribution of students by socioeconomic quintile is similar in Bridge and other private schools.

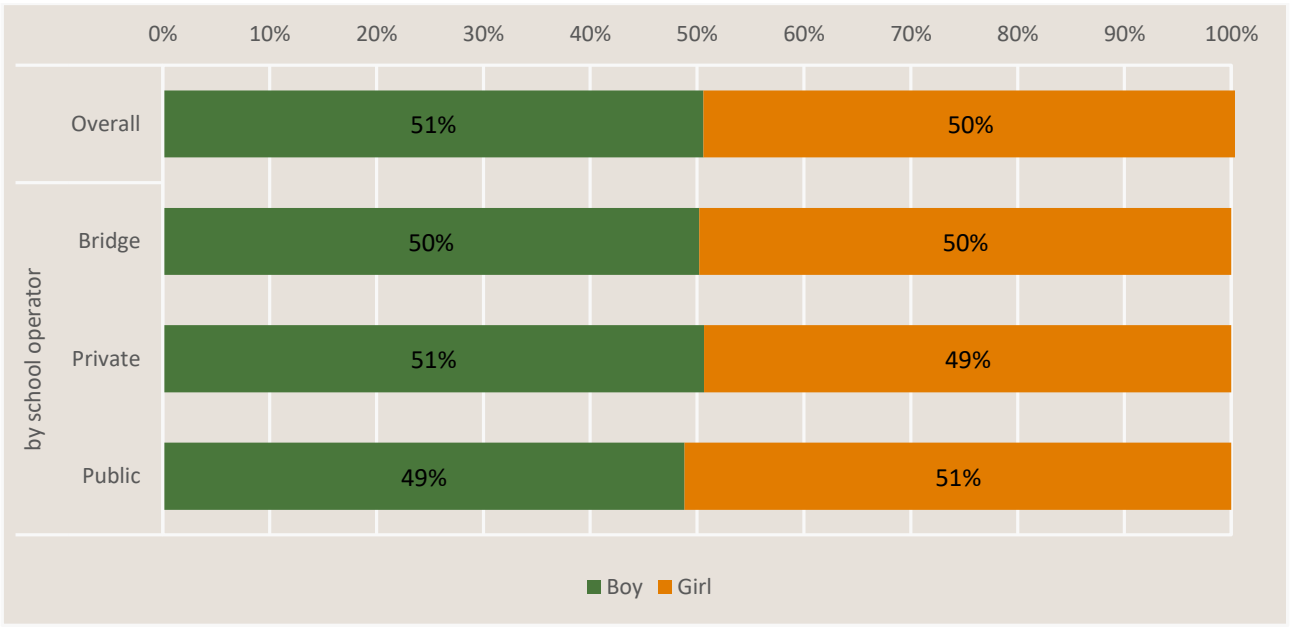
[11] The p-values of the differences between Bridge schools and public schools, and between private and public schools are 0.0000.

Figure 2 Distribution of socioeconomic quintile, by school type



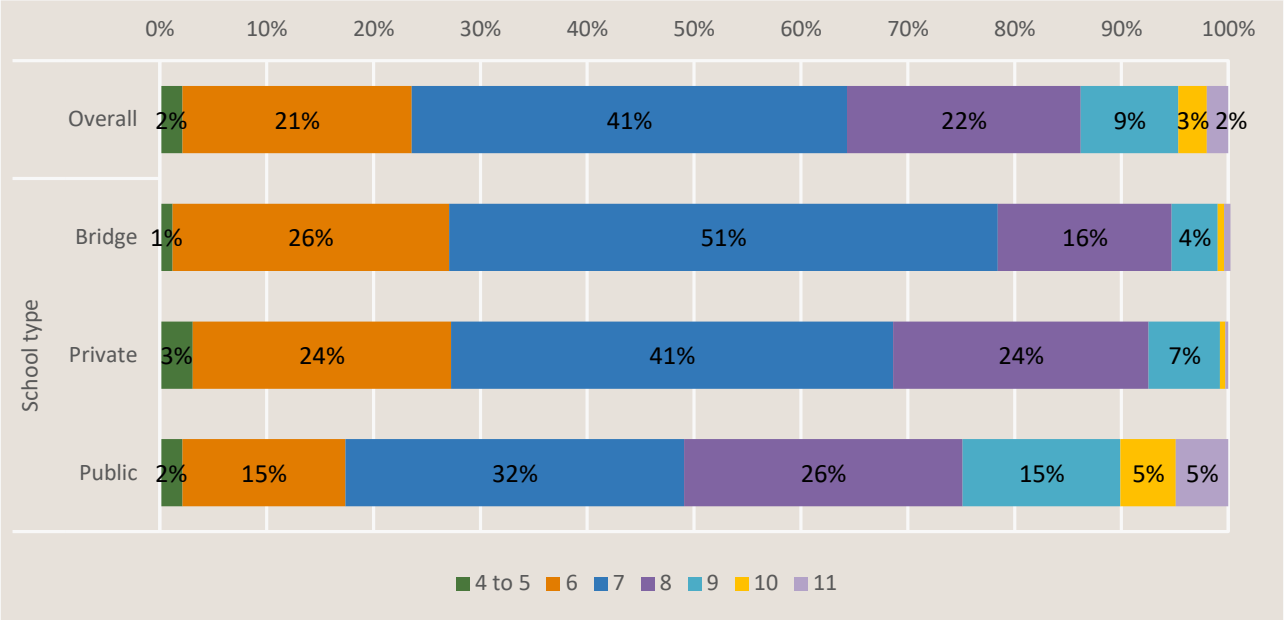
All of the schools surveyed have a balanced distribution of boys and girls.

Figure 3 Distribution of gender, by school type



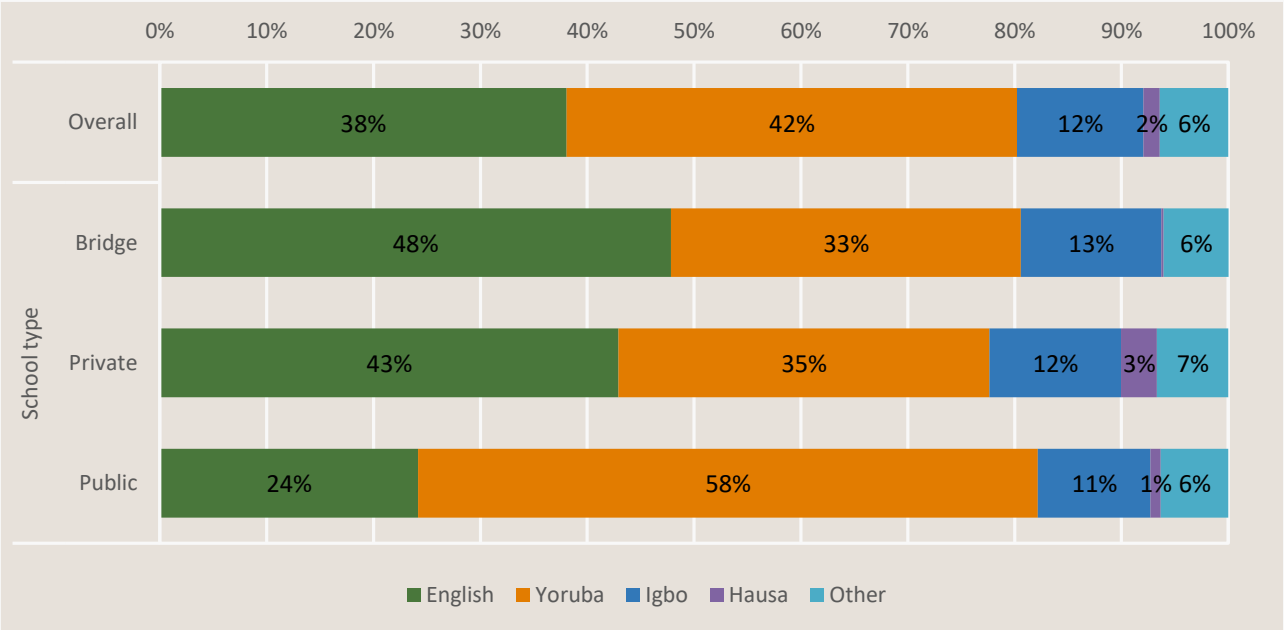
In Primary 2, children are expected to be between 6 and 8 years old: children in Bridge schools and other private schools are more likely to fall in this age category. 93% of children in Bridge schools, 89% of children in private schools, and 73% of children in public schools fall within this range, and the differences between schools are statistically significant[12].

Figure 4 Distribution of age, by school type



Children in Bridge schools and private schools are significantly more likely than children in public schools to primarily speak English at home. Yoruba is most frequently spoken at home amongst children in public schools.

Figure 5 Distribution of home language, by school type

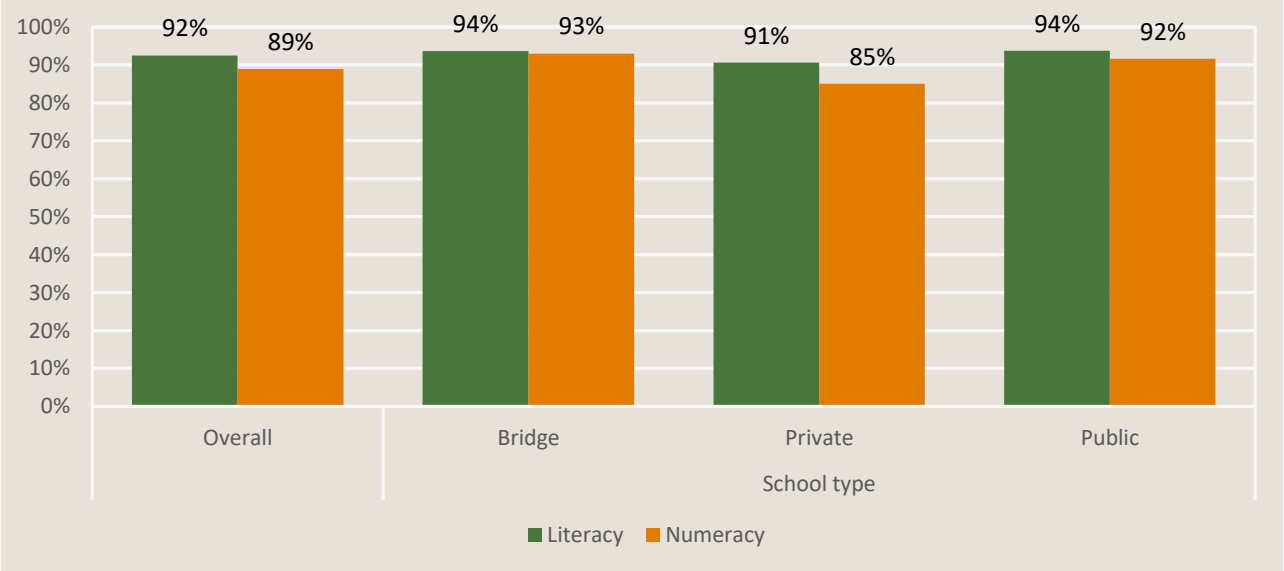


[12] The p-values of the differences between Bridge schools and other private schools, Bridge schools and public schools, and between private and public schools are 0.0000.

3.2.2 Teaching activity

Teachers in Bridge schools spend more time teaching numeracy than in other private schools (93% vs 85%), and a similar amount to public schools (93% vs 92%). Time spent teaching literacy is similar across all school types (94% in Bridge schools and public schools vs 91% in private schools). Teaching activities include teacher reading or lecturing pupils, leading group learning activities, writing on the blackboard, testing students or supervising and observing students while they read aloud, completed task or wrote on the board. Activities not counted as engaging in teaching include taking a break, not being in the classroom, and doing paperwork. It is important to note that since teachers are made aware that their classrooms are being observed, there is a possibility that teachers may deviate from their usual behaviour and engage more in teaching activities because of this research. There were no statistically significant differences in the time spent teaching for literacy classes while for numeracy classes, private school teachers spent 8% less time than Bridge teachers and 7% less time teaching than public teachers, and these differences are statistically significant.[13]

Figure 6 Average share of class time spent engaged in teaching activity



Researchers observing lessons judged that lessons in Bridge schools seemed as likely to be planned as lessons in private schools, and less likely to seem planned than in public schools. Overall, 75% of literacy and 73% of numeracy lessons observed seemed planned even though 82% of English teachers and 77% of mathematics teachers reported using a lesson plan (see figures below). Enumerators first asked teachers whether they were using a lesson plan - a list of activities to be carried out during the interval of the lesson and the time allocated to each of them, and only selected yes if the teachers were able to show them the lesson plan. Then, they were asked to note down whether the lesson seemed planned, and specifically whether the teacher was following a timed activity plan in the way they were teaching the lesson.

Public schools have the highest percentages of lessons that seemed planned and where the teacher was able to present a lesson plan to the enumerator, for both literacy and numeracy lessons.[14] This is a surprising finding, given that in Bridge schools all teachers are provided with pre-written detailed lesson plans, or teacher guides.

[13] P-values<0.05

[14] P-values<0.05

Figure 7 Percentage of literacy lessons that seemed planned/where teacher had a lesson plan

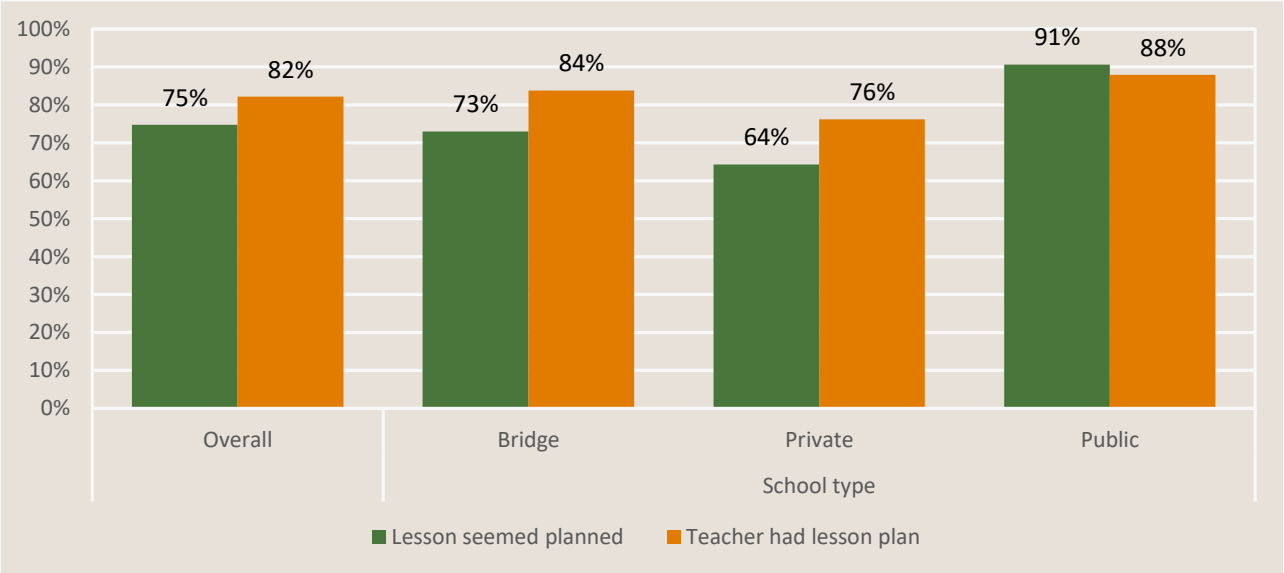
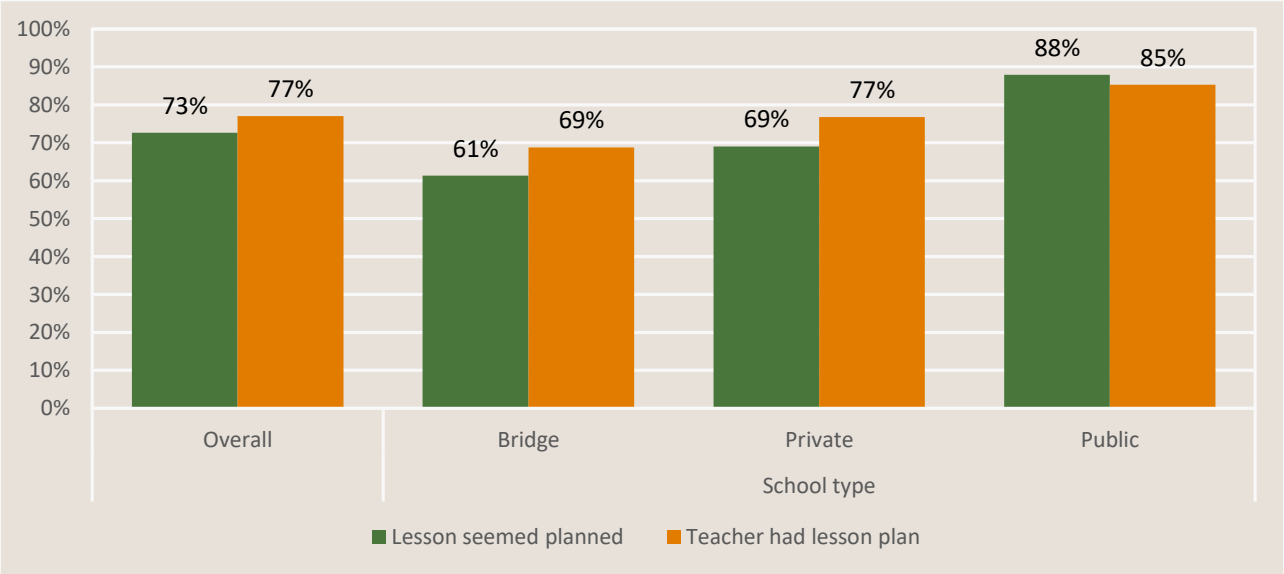


Figure 8 Percentage of numeracy lessons that seemed planned / where teacher has a lesson plan



Most time in all observed lessons was spent lecturing or presenting. Only a small proportion of classroom time is spent on asking questions from the teacher to a student in a whole-class setting. The distribution of the time is similar in the three school types, with teachers in Bridge schools most likely to be observed engaging in teaching activities, such as lecturing/presenting, testing/asking questions or supervising/leading pupils in activity[15].

[15] Note: the standard duration of a class is 30 minutes, so enumerators were instructed to remain in the classroom even if the teacher gives the students a break. A break is equivalent to time spent not teaching.

Figure 9 Mean proportion of time spent in each teaching activity during literacy class, by school type

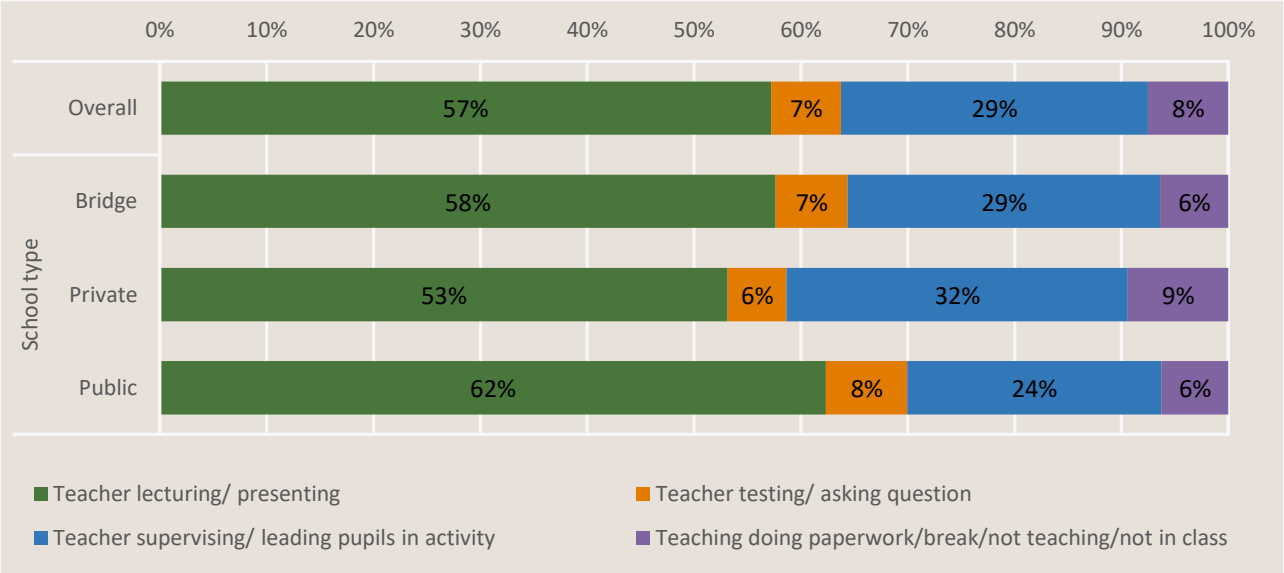
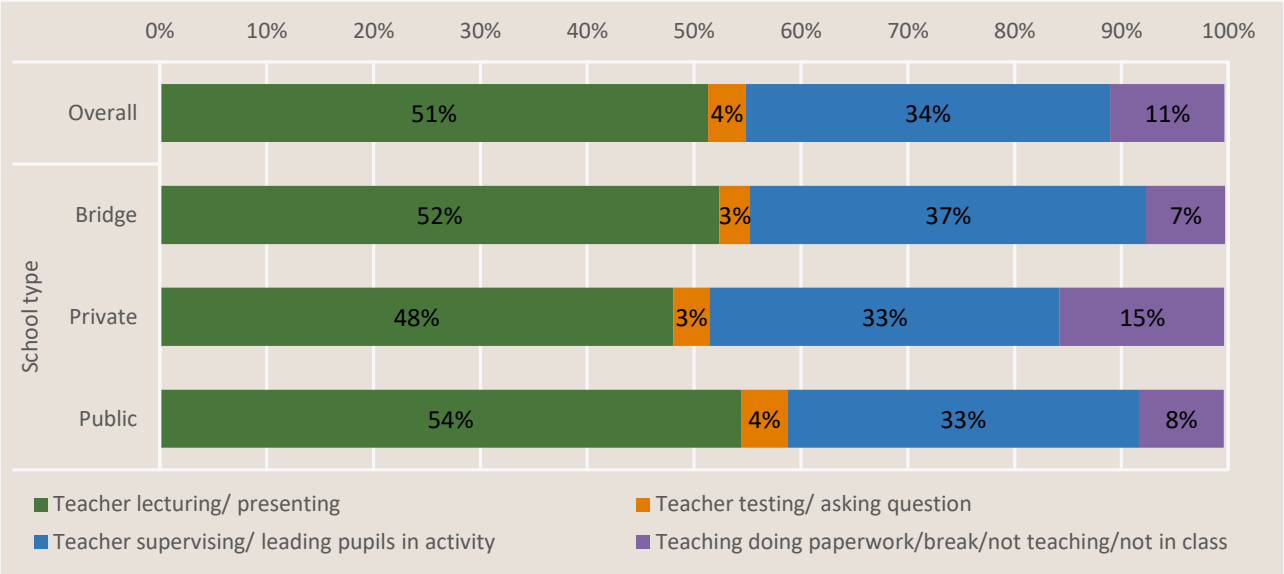


Figure 10 Mean proportion of time spent in each teaching activity during numeracy class, by school type

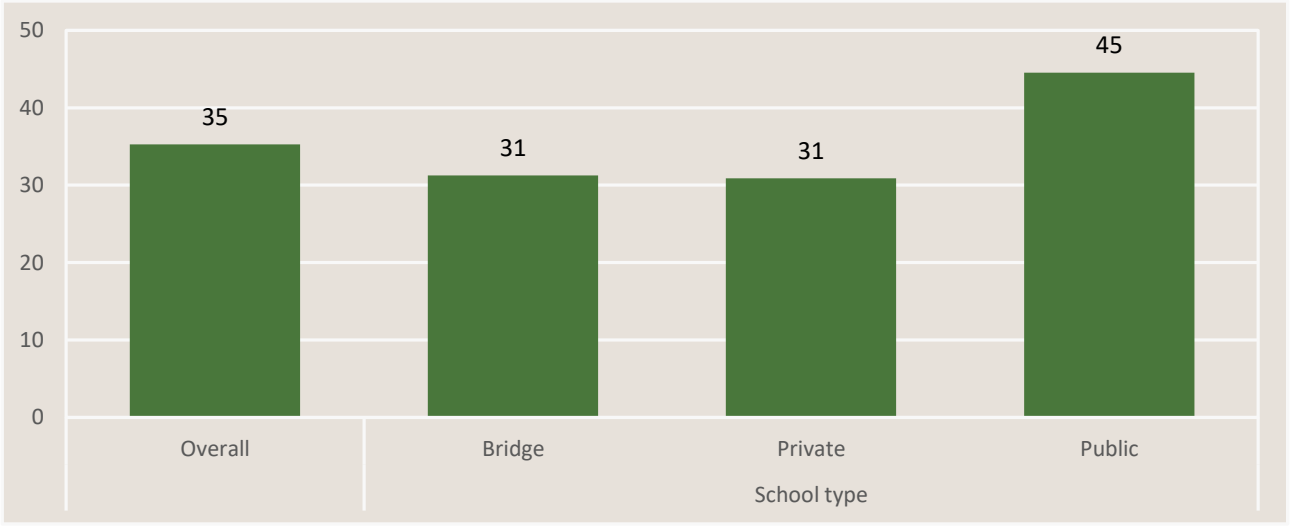


Teachers in Bridge schools were less likely to "hit, pinch, or slap" a child during a lesson (5%) than in private or public schools (both 31%).

3.2.3 Teacher characteristics

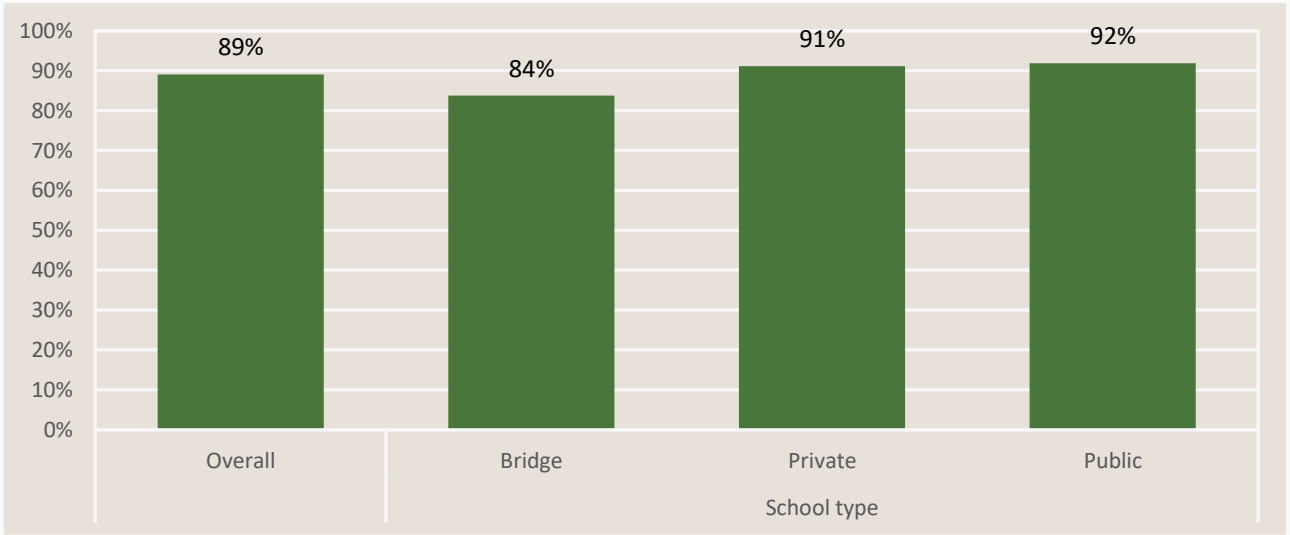
The average teacher in a Bridge school is 31 years old, similar to those in other private schools and significantly younger than those in public schools (45 years old).[16] While teachers at Bridge and private schools have similar ages, the average private school head teacher is nine years older than the average teacher (40 compared to 31), whereas at Bridge schools the average head teacher is only three years older than the average teacher (34 compared to 31).

Figure 11. Average (Mean) age of teacher



Most teachers are female (89%). Teachers in Bridge schools are slightly less likely to be female (84%) than teachers in other schools (91% in private and 92% in public), though the differences between groups are not statistically significant.

Figure 12 Share of female teachers



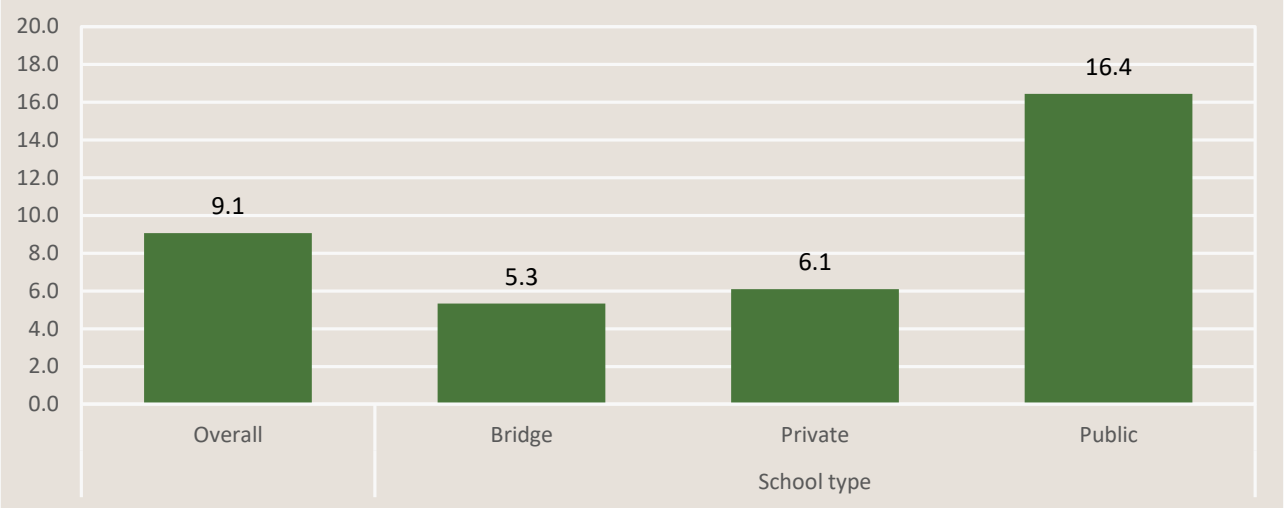
Teachers in Bridge schools have on average 5.3 years of experience, similar to those in private schools (6.1), and substantially less than those in public schools (16.4 years).[17] Despite this, Bridge and private

[16] P-value of Private/Public difference: 0.000; P-value of Public/Bridge difference: 0.000.

[17] P-value of Private/Public difference: 0.000; P-value of Public/Bridge difference: 0.000.

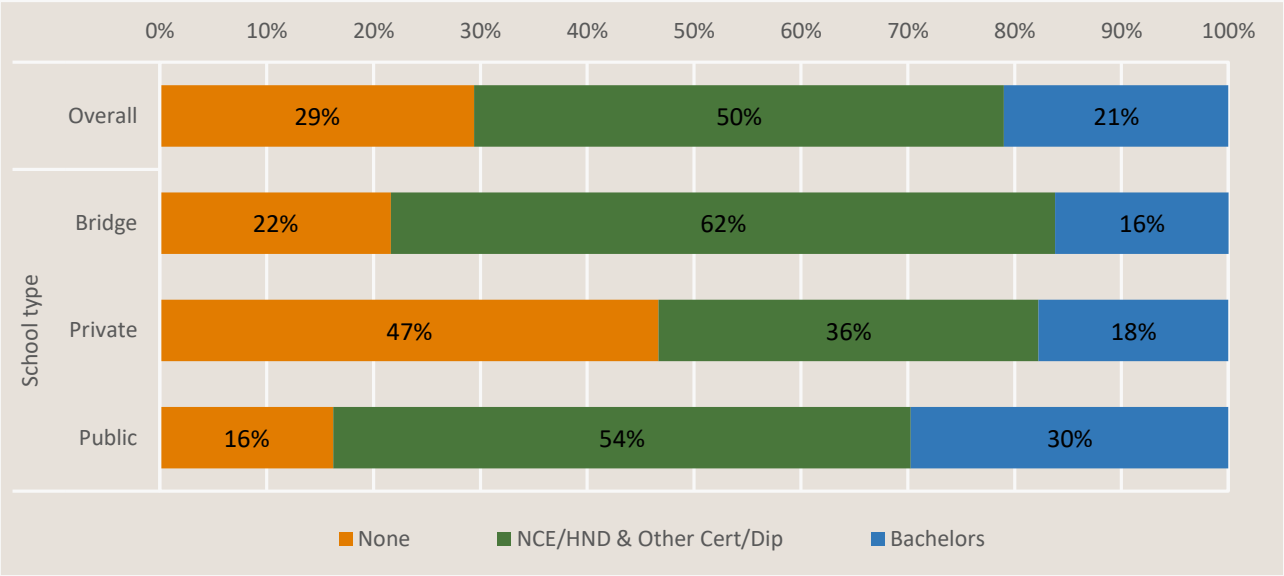
school teachers have more experience for their respective ages, as the gap in average age between public and non-public school teachers is 14, while the gap in the number of years of teaching experience is only 10-11 years.

Figure 13 Average number of years of teaching experience



Private schools have substantially more teachers without teaching qualifications. As illustrated in the figure below, 47% of teachers in private schools don't have teaching qualifications compared to 22% for Bridge and 16% in public schools. Public schools teachers are more likely to have a Bachelors (30%) than Bridge teachers (16%) and private school teachers (18%).

Figure 14 Breakdown of sampled teacher's level of teaching qualification (Bachelors including BEd and BA)



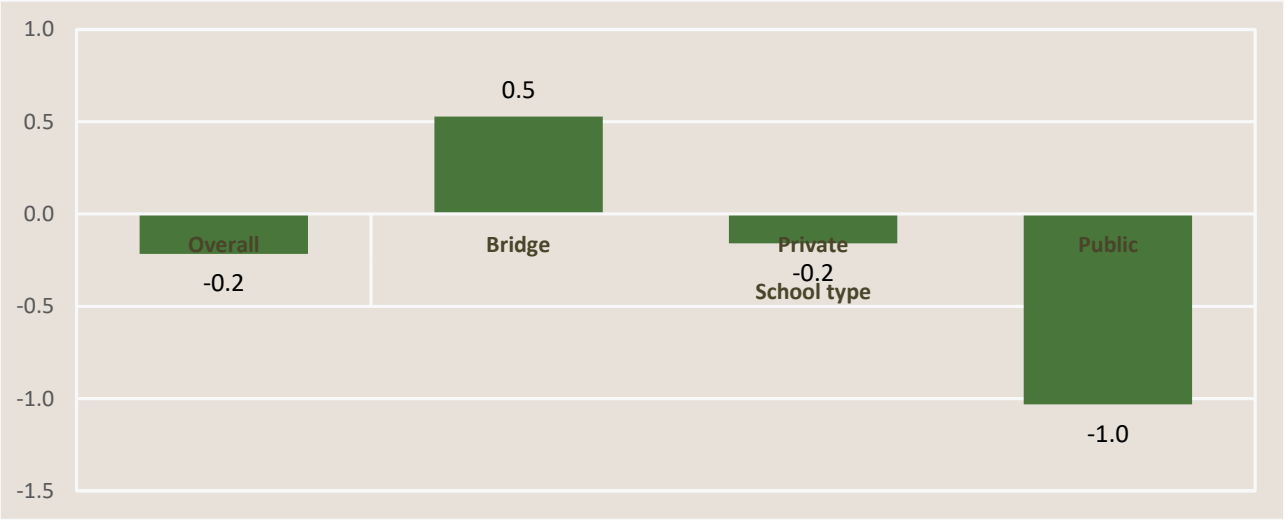
Note: NCE stands for National Certificate of Education (mandatory for teachers in public schools); HND stands for Higher National Diploma. Bachelors refers to a Bachelor's degree in Education or BA

3.2.4 Teacher motivation

Teacher motivation is highest for Bridge teachers, average for private schools and very low for public schools. The figure below presents teacher motivation scores ranging from -5 to 5 with a mean close to 0.

The difference between public and private schools is significant at the 10% level, whereas the difference between Bridge and public schools is significant at the 5% level.[18] Scores are based on an index of 29 questions about the teaching environment, career aspirations, teaching abilities and day-to-day enthusiasm for teaching.

Figure 15 Average teacher motivation score



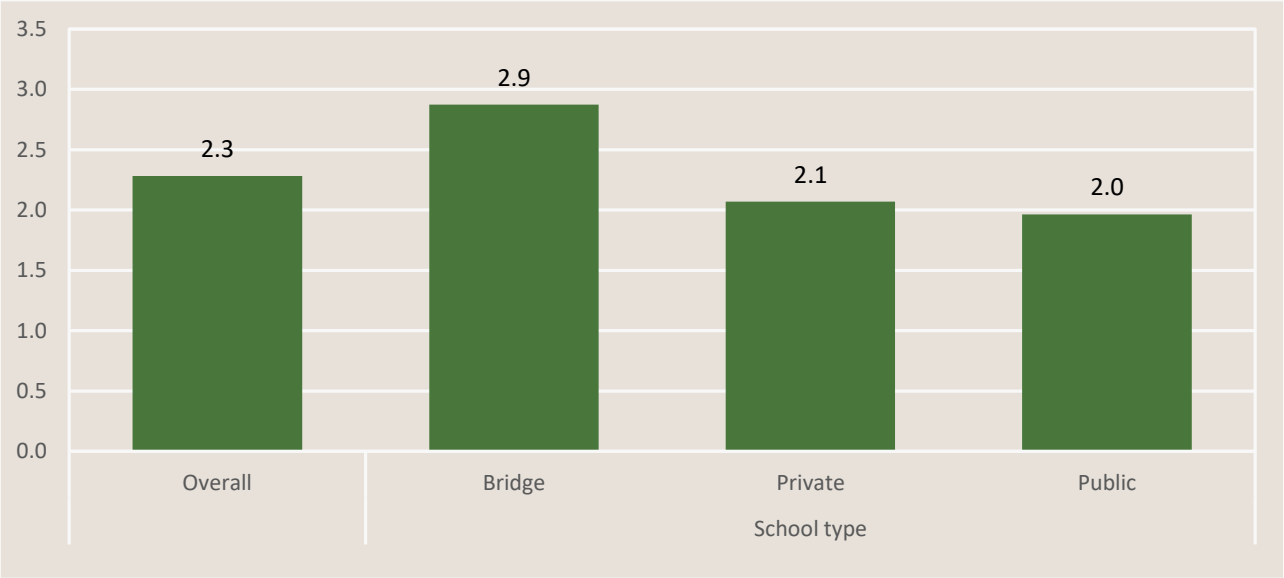
3.2.5 School management

Bridge schools have significantly higher management scores than private schools and public schools. The figure below shows the weighted mean scores using the four main component areas of management: operations, people management, monitoring, and target setting. Overall, the average score was 2.3, on a scale from 1 to 5. Bridge schools scored 2.9, compared to 2.1 for private schools and 2.0 for public schools.[19]

[18] P-value of Private/Public difference: 0.097; P-value of Public/ Bridge difference: 0.002.

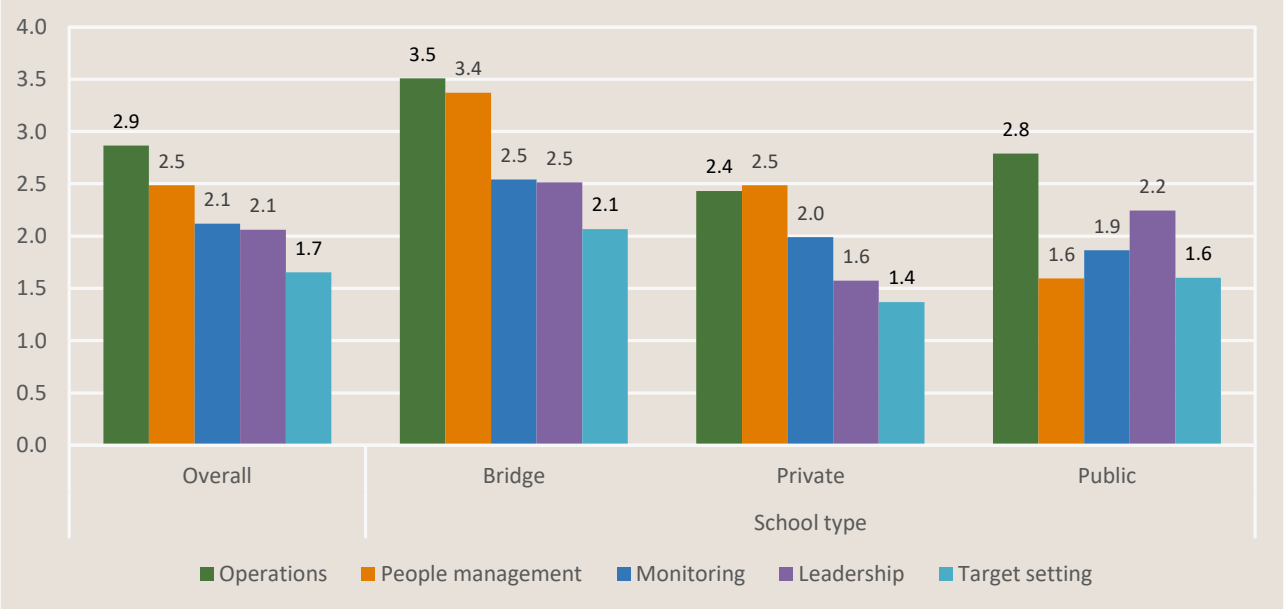
[19] P-value for Private/Bridge difference: 0.000; P-value for Public/Bridge difference: 0.000.

Figure 16 Overall average (mean) school management score



The high management scores in Bridge schools appear to be driven by above average performance in all sub-categories of management, and particularly operations and people management. Across school types, the lowest scores were for target setting, which suggests that few schools in Lagos link school management practices with robust measures of school quality such as learning outcomes.

Figure 17 Average school management score by sub-component



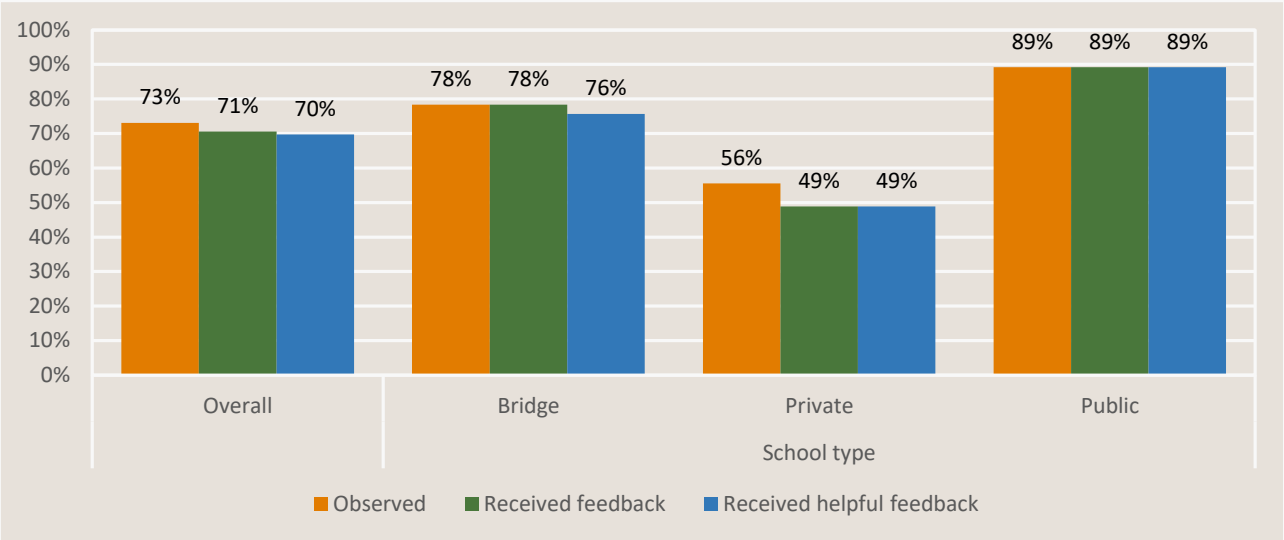
3.2.6 Monitoring (internal and external)

The lowest level of internal monitoring is in private schools. 56 percent of teachers in private schools had their lessons observed by a superior (e.g. head teacher, assistant head teacher or head of department). This is 22% points lower than Bridge and 43% points lower than public schools, with these differences being

statistically significant.[20] Academy Managers, Academic Field Officers and Teacher Trainers conduct regular lesson observations as part of the Bridge model and therefore, their rate of internal monitoring (78%) is similar to those of public schools (89%) with a difference that is not statistically significant.

Most teachers in Bridge and public schools both had their lessons observed and received feedback that they found to be helpful. Across all schools, out of 73% of teachers who report having their lessons observed, 71% report having received feedback, and 70% report receiving feedback that is helpful. In private schools, only 49% of teachers receive feedback despite 56% of teachers receiving monitoring visits.

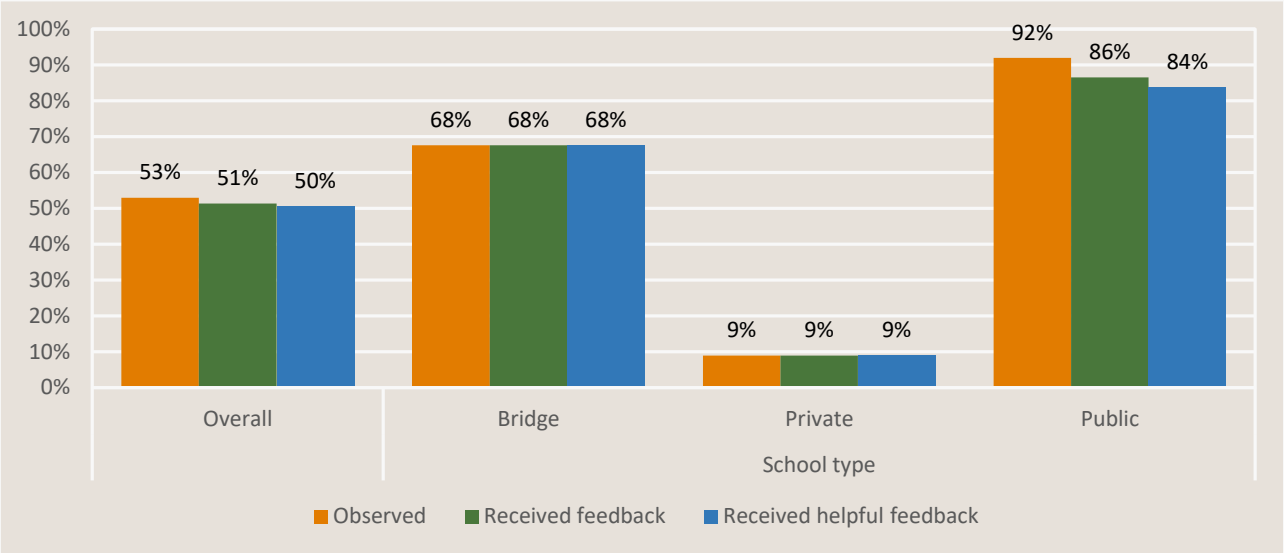
Figure 18 Prevalence of lesson observation and provision of feedback to teachers (internally)



Public schools have the highest level of external monitoring (e.g. from State Ministry of Education representatives or a Local Government Education Authority) through lesson observations. Of the teachers interviewed in public schools, 92% reported having received a visit from an external monitor, compared to 9% in private schools. For public schools the rate of external monitoring is higher than for internal monitoring. The opposite is the case for private schools, which have much less external monitoring compared to internal reviews, suggesting low levels of government monitoring of private schools. Bridge schools also report high levels of external monitoring, although lower than their internal monitoring levels. Whereas most internal monitoring is combined with helpful feedback in the view of public school teachers, not all external monitoring is combined with feedback, and not all feedback is perceived as helpful.

[20] P-value of Private/Bridge school difference: 0.025; P-value of Private/Public school difference: 0.000.

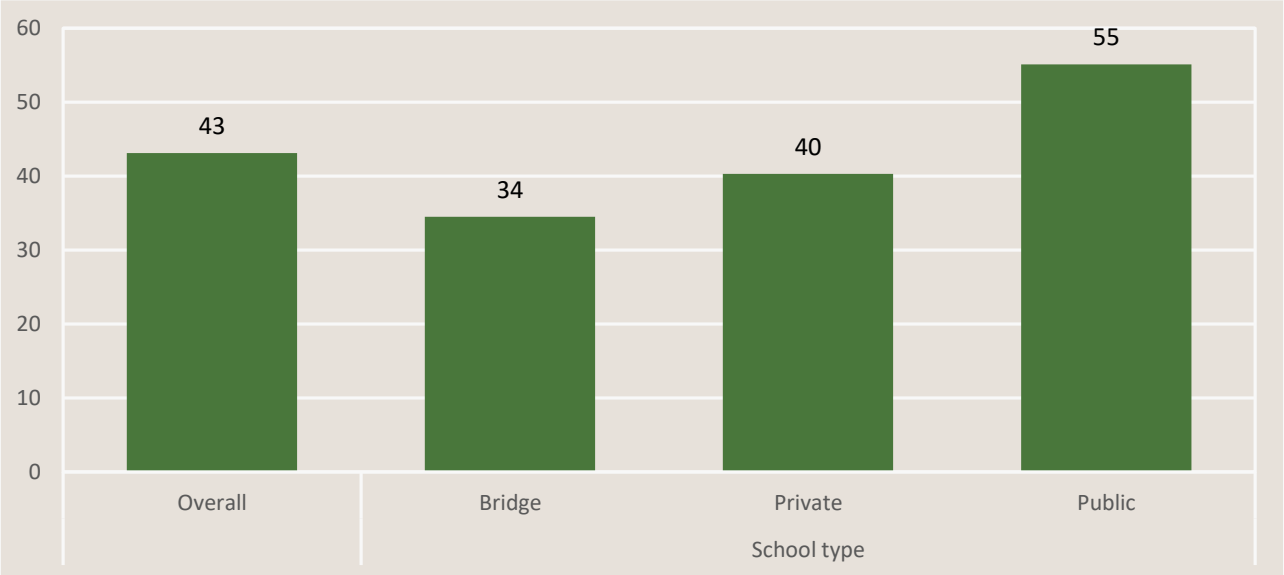
Figure 19 Prevalence of lesson observation and provision of feedback to teachers (externally)



3.2.7 Headteacher Characteristics

Bridge schools have the youngest head teachers, and public schools the oldest. While the overall mean age for head teachers in the sample is 43 years, public schools have significantly higher mean ages at 55 years, than private schools (40 years) and Bridge schools (34 years).[21] Mean head teacher ages in private schools are also significantly higher than at Bridge.[22]

Figure 20 Average (mean) age of head teacher



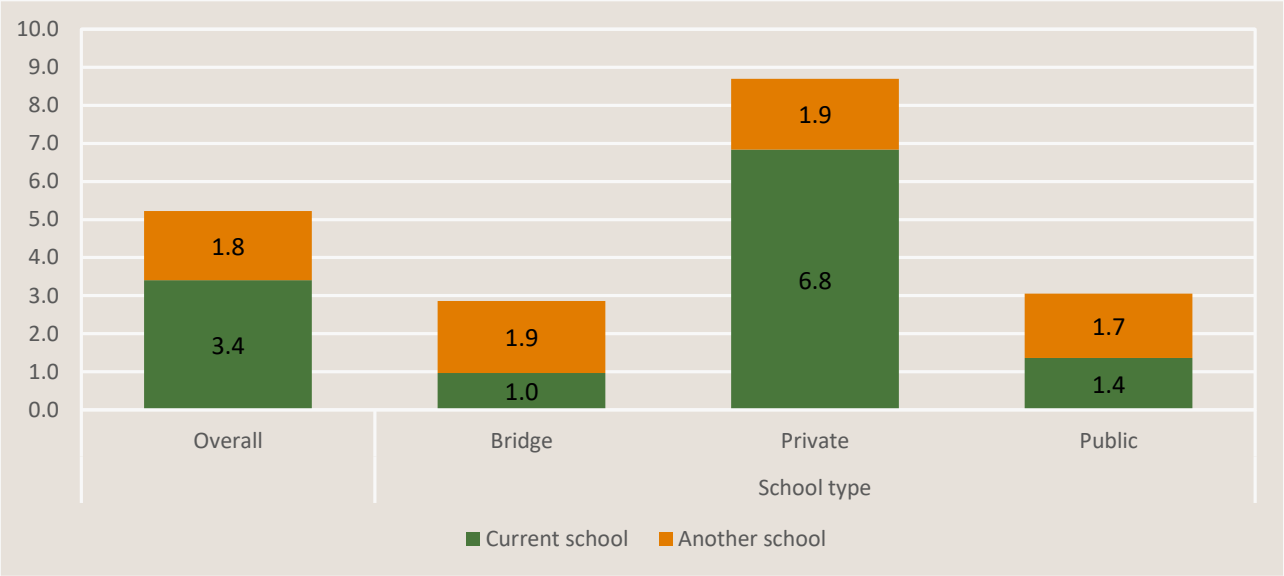
Private school head teachers have substantially more school management experience than teachers in Bridge and public schools. As illustrated in the figure below, head teachers have an average of 5.2 years of experience in a school management position, with an average of 3.4 years at their current school and 1.8 at another schools. Private school head teachers have a substantial and statistically significant amount more

[21] P-value of Public/Bridge school difference: 0.000; P-value of Public/Private school difference: 0.000.

[22] P-value of Bridge/Private school difference: 0.001.

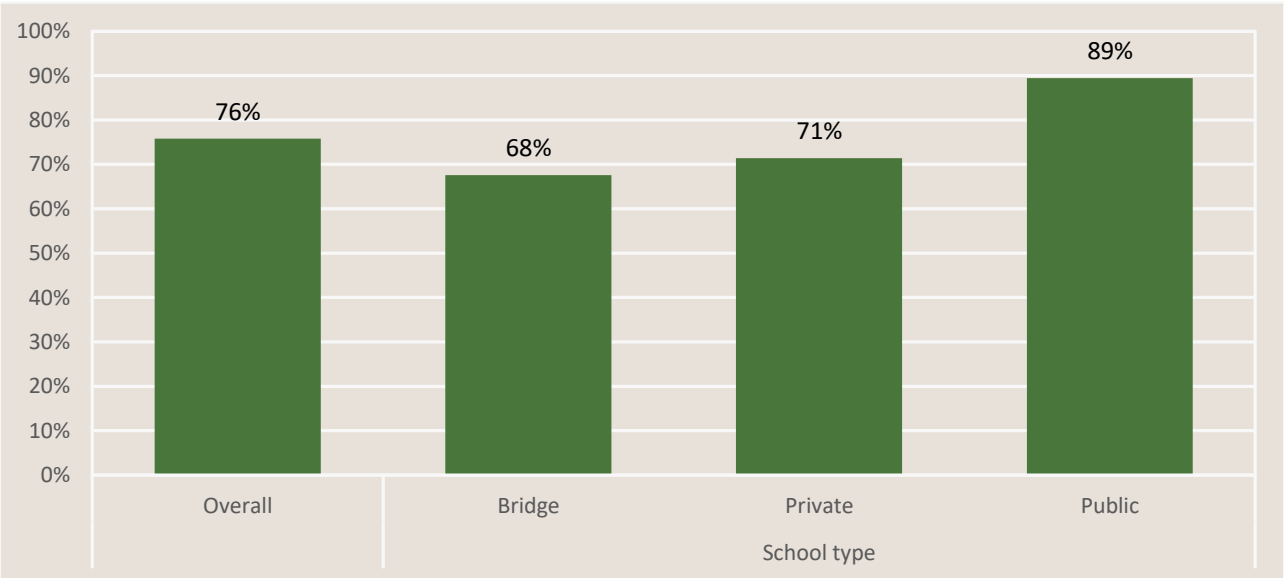
years of experience than Bridge and public schools. Given age differences by school type, the trends suggest that in public schools, educators became head teachers later in their careers. The fact that 74% of head teachers have management experience at a different school suggests that high head teacher turnover is a possibility.

Figure 21 Average head teacher management experience (top part of chart indicates years at current school, bottom part indicates years at a different school)



Head teachers in public schools are more likely to be female than head teachers in private schools and Bridge schools. As illustrated in the figure below, 76% of all head teachers are female. This percentage is significantly higher in public schools (89%), than Bridge and private schools (68% and 71%, respectively).[23]

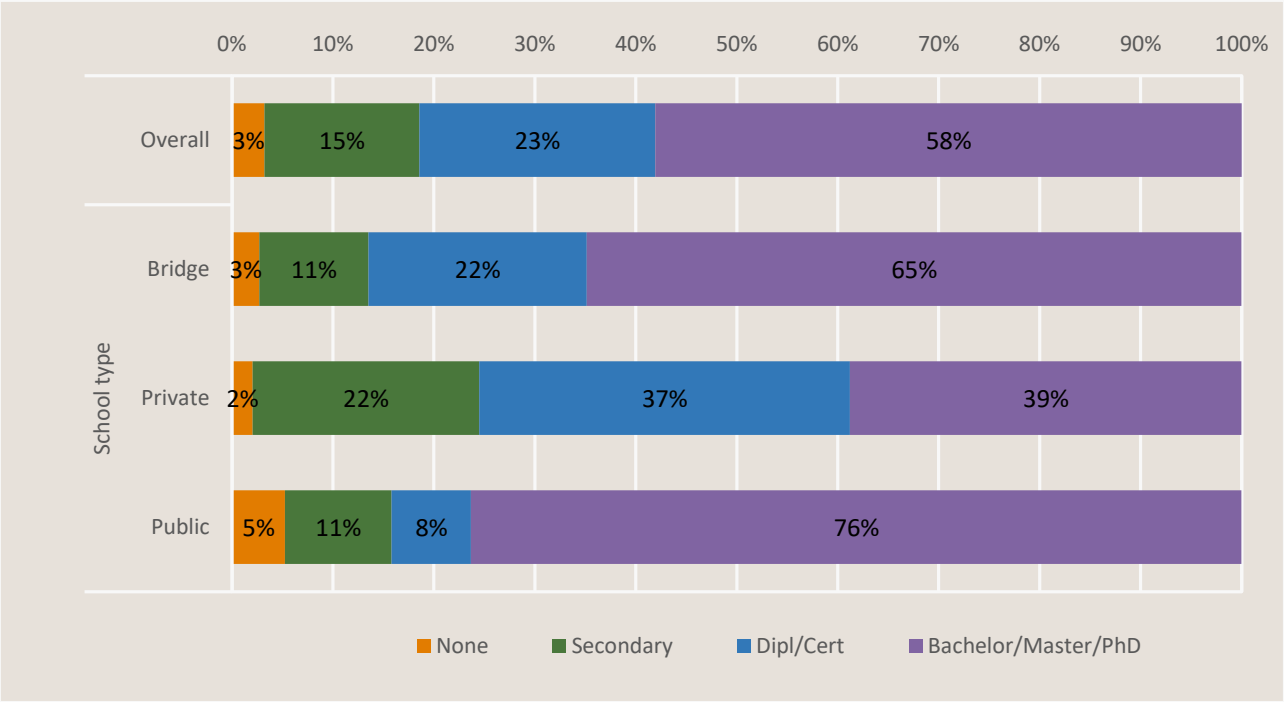
Figure 22 Share of female head teachers



[23] P-value of Public/Bridge school difference: 0.019; P-value of Public/Private school difference: 0.029.

Head teachers in public schools and at Bridge have the highest level of education. As illustrated in the figure below, 76% of public and 65% of Bridge head teachers have university degrees. This is the case for only 39% of private school teachers.

Figure 23 Head teacher's level of completed education



3.2.8 School fees

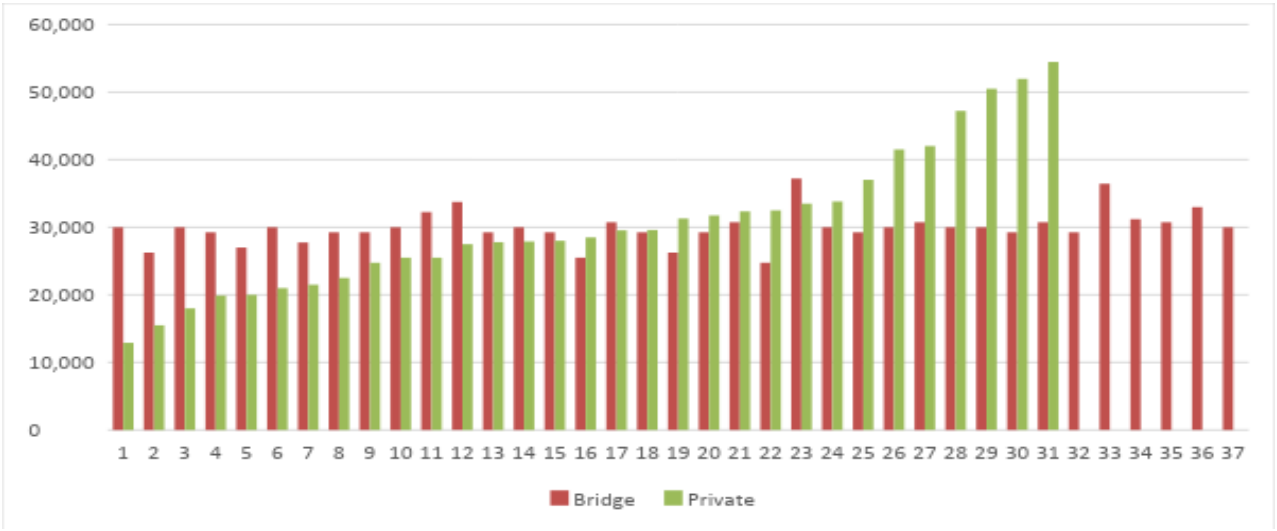
School fees reported by the average parent are consistently higher than school fees reported by the head teacher, although the difference is small. Overall, the mean school fees per year according to the head teacher is 20,708 Naira ($57) and 22,803 ($63) according to the average parent, and the difference is statistically significant[24]. While average private school fees are not significantly higher than Bridge schools, their variance is greater. This is due to the fact that Bridge schools follow a standard fee structure, resulting in a fee range between 24,750 ($69) and 37,200 Naira ($103), while the range for private schools is 12,900 ($36) to 62,000 Naira ($172) (according to the head teacher).

Table 5 Average school fees as reported by head teacher and parents, by school type

Average total fees per year, in Naira	Overall	Bridge	Private	Public
As reported by head teacher	20,708 ($57)	29, 936 ($83)	32,072 ($89)	Data not collected*
Std error	401	104	533	
As reported by parents	22,803 ($63)	30,474 ($85)	35,711 ($99)	994 ($3)
Std error	444	232	660	93

Bridge schools and the private schools with which they have been matched have broadly similar fee levels. The figure below shows side-by-side comparisons of the schools fees of Bridge schools and their private school matches, illustrating relatively similar fee levels for matched schools.[25]

Figure 24 Comparison of school fees between Bridge schools and their private school matches



[24] P-value 0.0000; P-value of the difference in Bridge schools is 0.0378; P-value of the difference in private schools is smaller than 0.0001 P-value 0.0000

[25] Some of the Bridge schools had two matched schools in which case the average of the two school fees was used. Note that this juxtaposition of fee levels across schools does not take into account the fact that the matching was based on a distance variable consisting of three separate indicators, including school fees.

Note: Five Bridge do not have a match in the sample due to private school refusals to participate in the survey, discussed in section 2.3. Where a Bridge school has two matches, the average of fees between the two is represented in the figure.

School fee levels (used in analysis)

Most schools in our sample are low cost, and seven are medium cost. The table below displays the distribution of school fees according to the delineations in Tooley & Yngstrom (2014), and using the data as reported by the head teacher. We classify public schools in a separate "no fee" category for the analysis, while recognising that parents do report spending a small amount of money (under 1,000 Naira per year, i.e. approximately $2.75) on public schools.

Table 6 Distribution of school fee levels

	Bridge	Private	Public	Total
No fee (0)	0	0	38	**38**
Low (1-42,802)	37	37	0	**74**
Medium (42,802-85,605)	0	7	0	**7**
High (>85,605)	0	0	0	**0**
Total	**37**	**44**	**38**	**119**

3.3 Correlations between student, household, teacher, & school characteristics with learning

To help anchor the comparisons in this section, the table below presents the mean, minimum and maximum literacy and numeracy scale scores by school type.

Table 7 Literacy and numeracy scale scores by school type

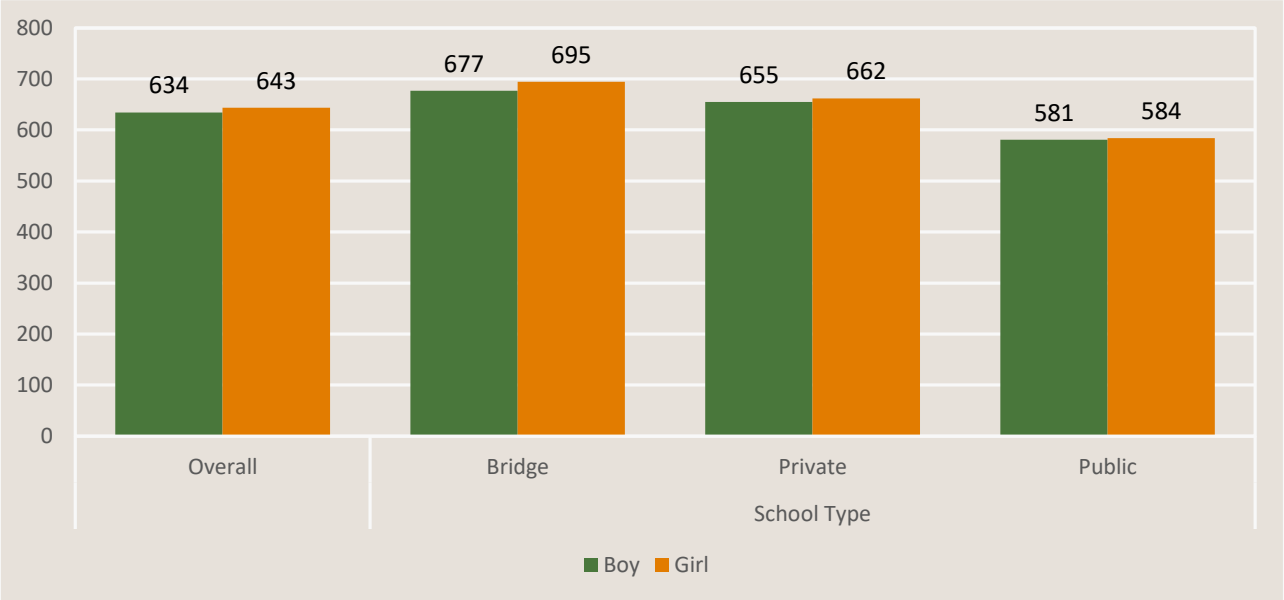
	Overall			Bridge			Private			Public		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
Literacy scale score	436	645	800	520	686	800	468	658	800	436	582	800
Std. error		2.1			2.6			3.4			2.9	
Numeracy	306	566	793	442	580	793	351	583	758	306	531	694
Std. error		1.6			2.0			2.8			2.7	

3.3.1 Student characteristics

Girls have higher literacy scores than boys across all school types (this differences is only statistically significant for students in Bridge schools and public schools).[26]

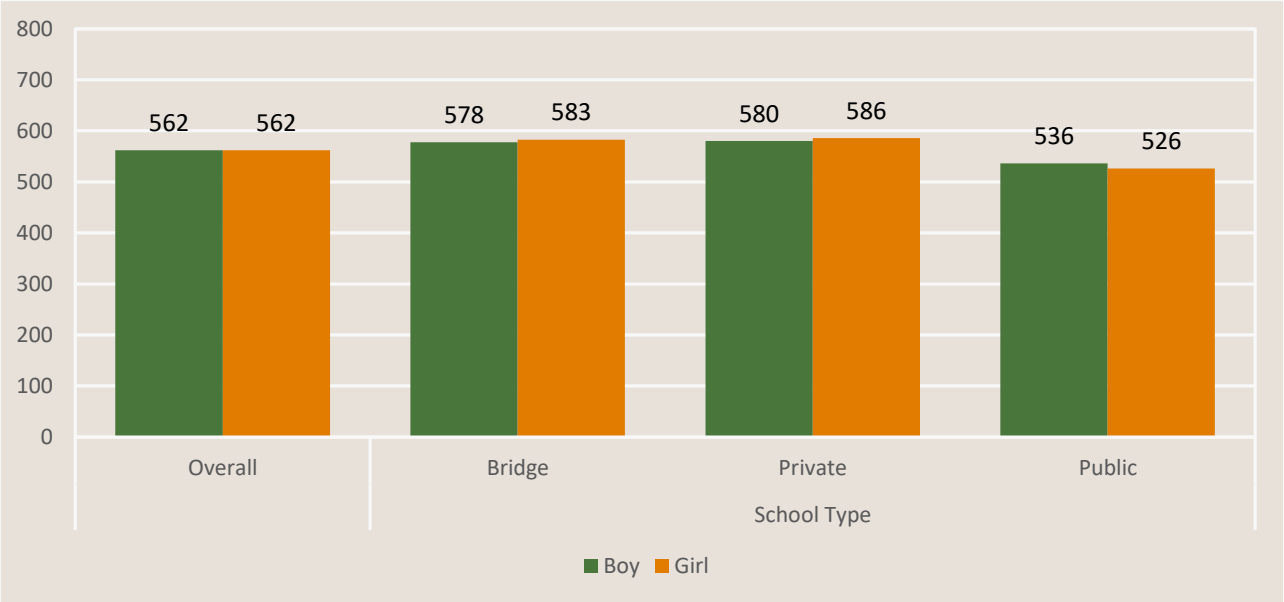
Numeracy performance across school types is similar between boys and girls. These findings are consistent with previous evidence. A study in Lagos at the Primary 4 level has found a statistically significant disadvantage for boys in literacy in public schools (EDOREN, 2017) and no significant gender differences in literacy for boys at the Primary 2 level in private schools (EDOREN, 2016).

Figure 25 Mean literacy scale scores by gender and school type



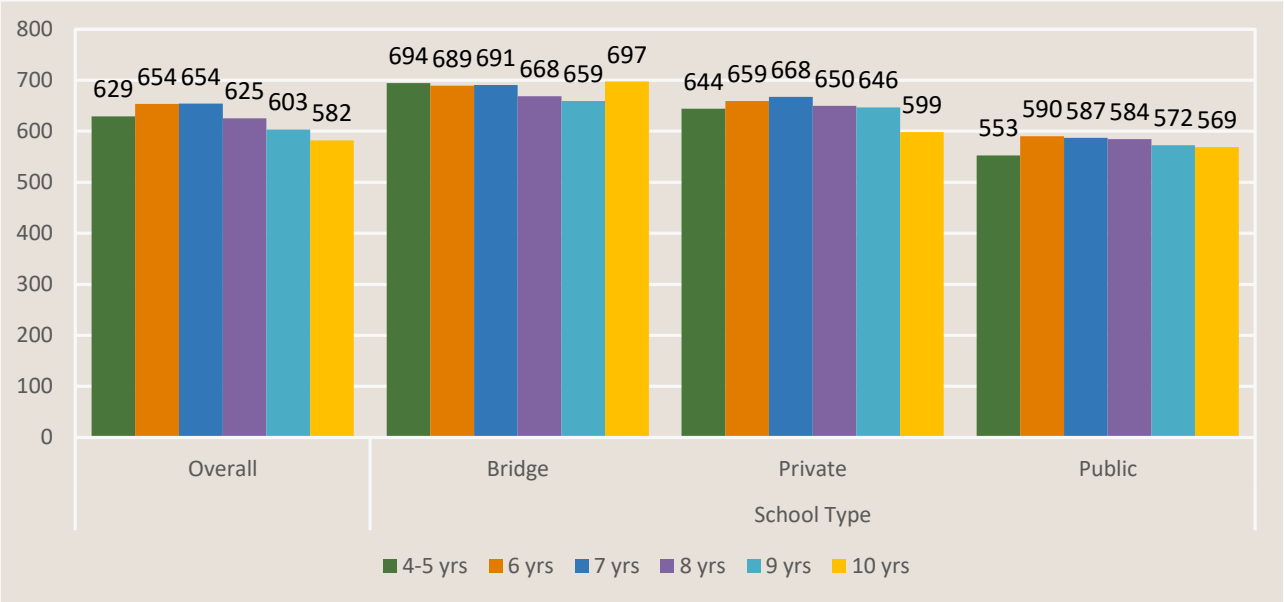
[26] P-value of 0.016 and 0.003, respectively

Figure 26 Mean numeracy scores by gender and school type



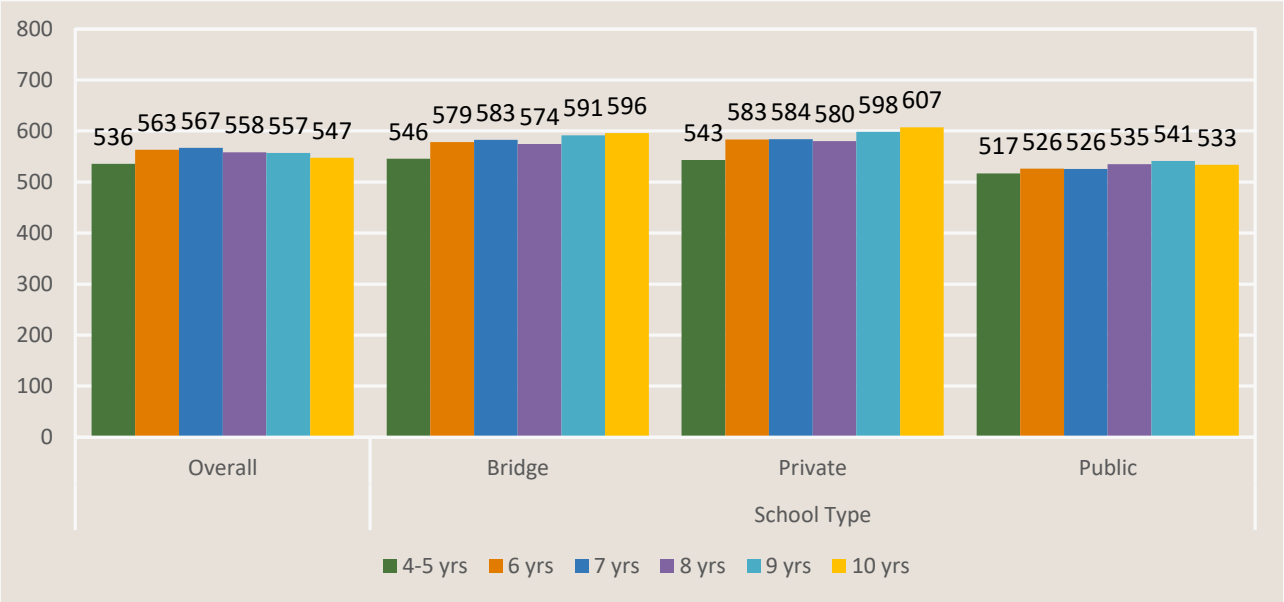
There is a negative relationship between being overage and achievement, which is stronger for literacy than numeracy. The figures below illustrate that for both literacy and numeracy, scores peak at the appropriate Primary 2 ages of 6 and 7 years[27]. The overall numeracy trends are similar to overall literacy trends. However when examining trends by school type, the trend for numeracy is more linear than the trend for literacy.

Figure 27 Mean literacy scores by age and school type



[27] The difference in literacy and numeracy scores between those older than 7 and those aged 7 and younger is negative and statistically significant with a p-value of 0.0000

Figure 28 Mean numeracy scores by age and school type



3.3.2 Household characteristics

Poorer students in our sample tend to have lower literacy scores and numeracy scores. In the following analysis we sort students into five quintiles based on their family score on a standard asset index questionnaire. The bars on the left of the figures below show a steady increase of literacy and numeracy scores from quintile 1 (poorest) up to the fourth quintile with statistically significant differences between adjacent quintiles and then a drop between the fourth and fifth.

Figure 29 Mean literacy scores by socioeconomic quintile and school type

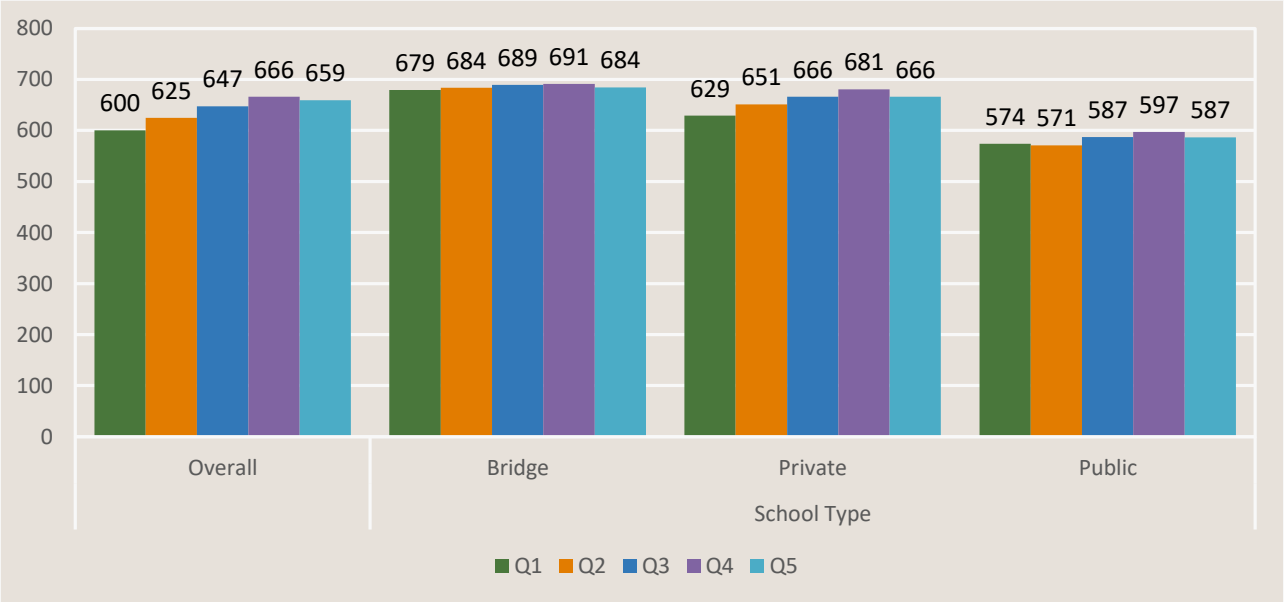
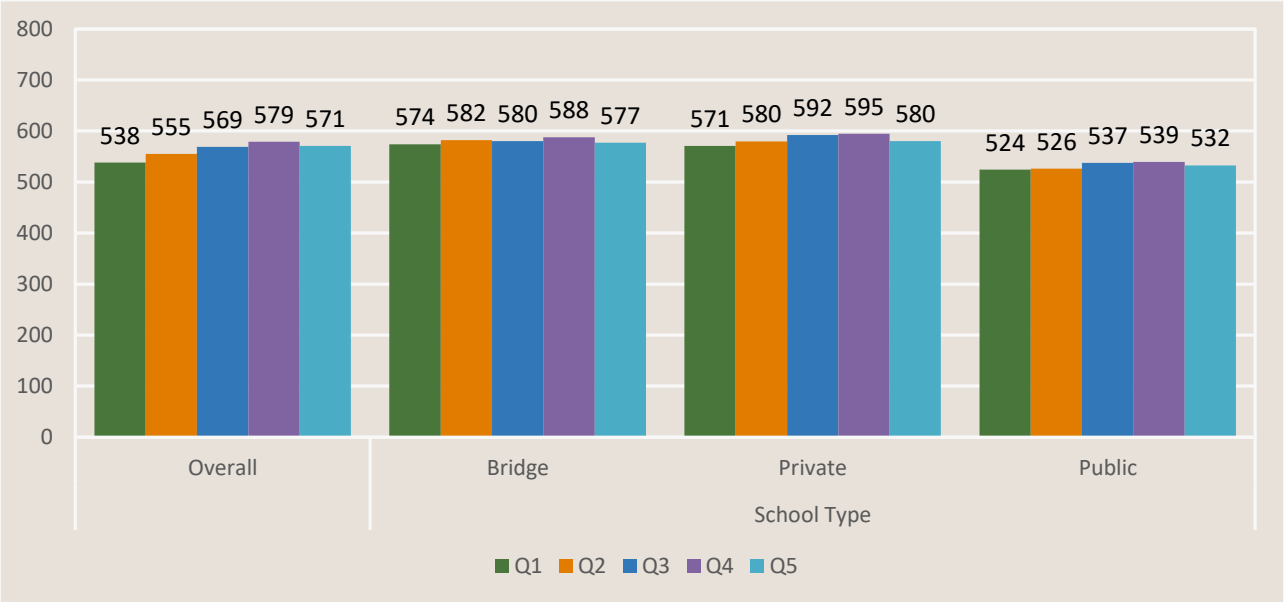
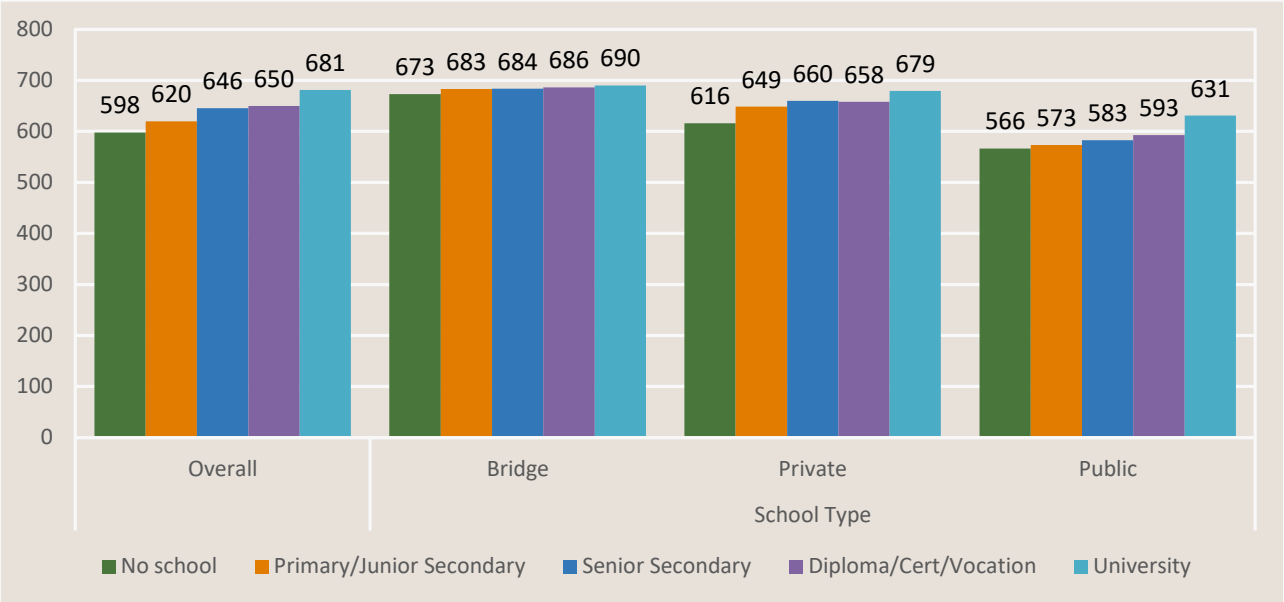


Figure 30 Mean numeracy scores by socioeconomic quintile and school type



Parents' education is positively associated with test scores (both literacy and numeracy) in public and private schools, but not at Bridge. Students in private and public schools show statistically significant associations between parent education levels and literacy/numeracy scores[28]. The same is not true of Bridge schools, where we do not find a strong relationship between parent education and student test scores.[29]

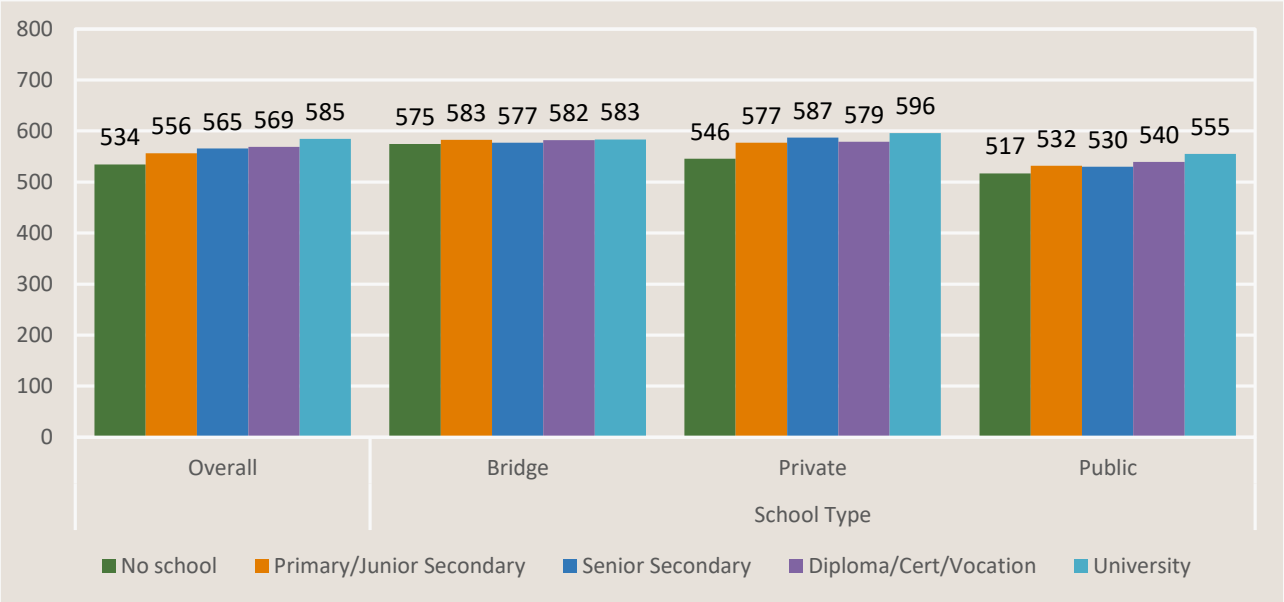
Figure 31 Mean literacy scores by parental education level and school type



[28] P-value for positive differences between adjacent education levels for literacy/numeracy scores of private and public pupils is <0.1.

[29] P-value for difference between no education and university education for literacy scores of Bridge pupils: 0.270; p-value for numeracy scores: 0.420.

Figure 32 Mean numeracy scores by parental education level and school type



Speaking Yoruba at home is correlated with poorer literacy outcomes. Primary home language is not associated with numeracy achievement. As shown below, overall children who speak Yoruba at home have lower scores than those speaking other languages, and the difference is statistically significant. There are no marked differences in numeracy performance by language spoken at home.

Figure 33 Mean literacy scores by home language and school type

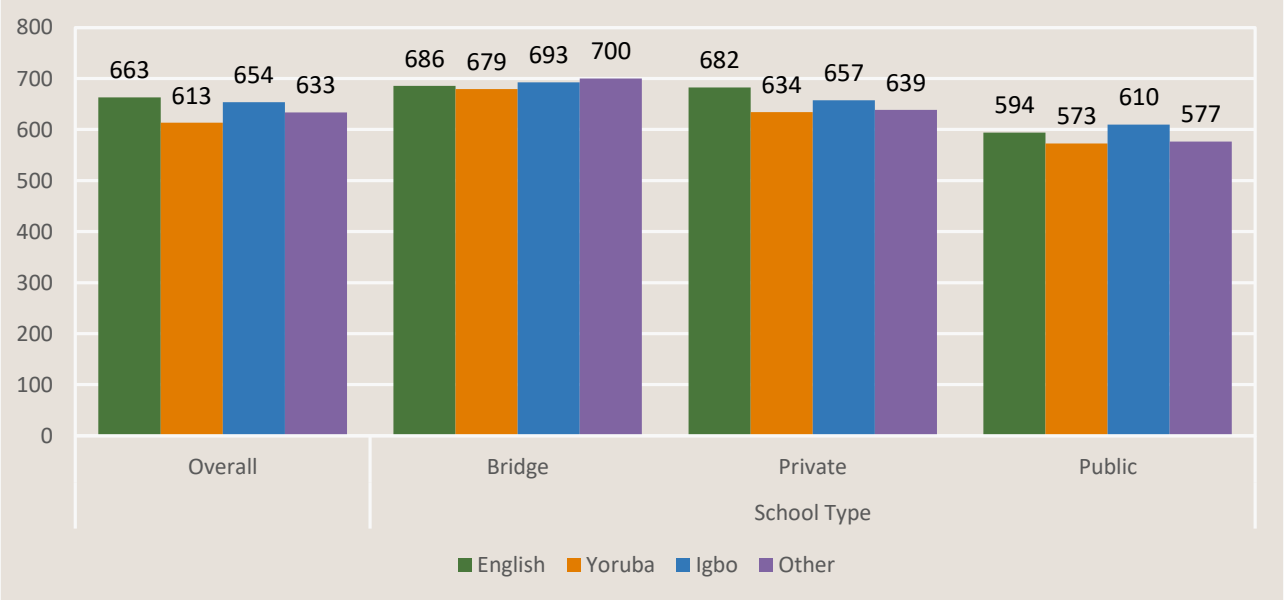
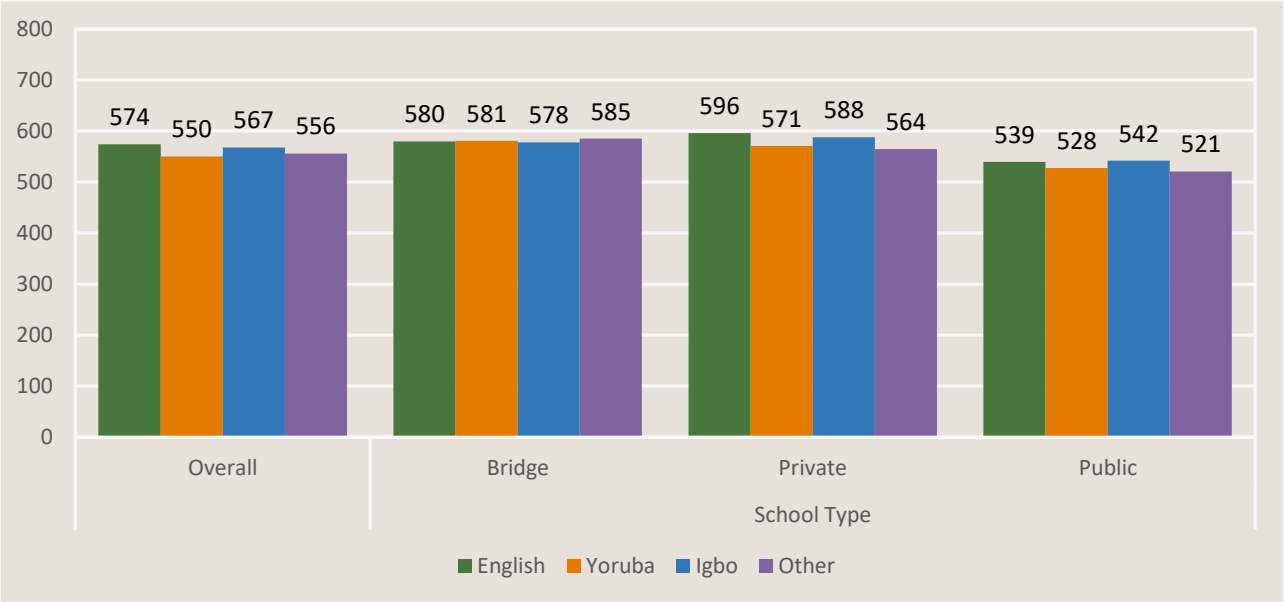


Figure 34 Mean numeracy scores by home language and school type



3.4 OLS Regressions

Regression estimates are obtained applying pupil sampling weights, with robust standard errors clustered by school. The dependent variables are literacy and numeracy achievement scores standardised around a mean of 0. In addition to the characteristics explored in the previous section, the regressions include local government area (LGA) fixed effects, which controls for all LGA observable and unobservable characteristics. We use regressions to understand how differences within school types correlate with levels of achievement. The regression results are reported in Tables 7 and 8 below.

3.4.1 Pupil and household characteristics

After accounting for other factors, wealthier students have better literacy scores but not better numeracy scores. This overall pattern varies by school type. At non-Bridge schools, wealth matters for literacy. Across school types, wealth is not associated with numeracy performance.

The correlation between student age and test scores is not statistically significant, after controlling for other factors.

Girls perform better than boys in literacy after controlling for other factors in Bridge schools (0.18 standard deviations), and public schools (0.11 standard deviations), but no better in numeracy. The gender gap in literacy is not statistically significant when looking only at other private schools.

Speaking English at home is associated with higher achievement in literacy in private schools (0.24 standard deviations). There are no statistically significant associations between the language spoken at home and numeracy achievement.

3.4.2 Teacher Characteristics and Student Learning

Whether the lesson seemed planned does not appear to be consistently correlated with learning, after controlling other factors. There is a strong positive correlation between a perceived plan and numeracy and both literacy and numeracy achievement in private schools. The percentage of time the teacher was engaged in teaching activity, while insignificant across school types for literacy, is negatively associated with numeracy scores (although the magnitudes of the coefficients are very small).

We do not observe a consistent relationship between teaching degrees and learning achievement. Neither teaching degrees (i.e. Nigeria Certificate for Education, Diploma Certificate for Education, etc.) nor Bachelor's degrees appear to be significantly correlated with learning.

We do not observe a strong relationship between teacher motivation and achievement. The only associations in the data are negative (for literacy in public schools and for numeracy in private schools).

3.4.3 School management and Student Learning

We observe a strong and statistically significant relationship between school management practice and literacy and numeracy achievement. This relationship is weakest (and not statistically significant) in non-Bridge private schools. A 1-point increase in the school management score (e.g. the transition between no evidence of a good practice to some evidence of a good practice) is associated with 0.19 standard deviation higher literacy scores in Bridge schools, and 0.30 standard deviations higher literacy scores in public schools. In Bridge schools, better management is correlated with numeracy scores which are 0.29 standard deviations higher.

There is no consistent relationship between either the amount of the head teacher's management experience or their qualifications and student learning, after controlling for other factors. Management experience is positively correlated with student literacy at private schools (significantly for non-Bridge schools), and negatively correlated with literacy in public schools. For numeracy, management experience is negatively correlated with achievement in Bridge schools and public schools, and positively and significantly correlated with achievement in private schools.

Table 8 OLS Regressions of literacy score factors

	(1) Bridge	(2) Private	(3) Public
Pupil is male	-0.182**	-0.087	-0.113*
	(0.069)	(0.106)	(0.061)
Age	0.364	0.349	0.288
	(0.339)	(0.280)	(0.202)
Age squared	-0.032	-0.021	-0.019
	(0.023)	(0.018)	(0.012)
Home language is English	-0.140	0.239**	0.183
	(0.106)	(0.101)	(0.111)
Home language is Yoruba	-0.164	-0.192	-0.128
	(0.118)	(0.119)	(0.104)
Continuous household wealth index	0.013	0.082**	0.080**
	(0.029)	(0.035)	(0.037)
Head teacher total years of management experience	0.004	0.059***	-0.012
	(0.009)	(0.021)	(0.013)
Head teacher has university degree	-0.019	0.122	-0.092
	(0.082)	(0.229)	(0.120)
School management average score 1-5 (avg of 15 questions)	0.189**	0.307	0.302***
	(0.084)	(0.403)	(0.100)
School fees according head teacher	0.027	-0.003	
	(0.026)	(0.007)	
Teacher has a teaching qualification (no bachelors)	0.166	0.034	-0.192
	(0.100)	(0.251)	(0.138)
Teacher has a bachelors' degree	-0.124	0.042	0.010
	(0.112)	(0.182)	(0.129)
Teacher motivation index (index of 29 Qs)	-0.035	-0.121**	-0.003
	(0.024)	(0.045)	(0.023)
Literacy lesson seemed planned	-0.098	0.439**	-0.272***
	(0.097)	(0.183)	(0.091)
Percentage of classroom time spent on teaching activity	-0.010**	0.003	-0.001
	(0.004)	(0.006)	(0.003)
Constant	-0.638	-2.932*	-1.922**
	(1.599)	(1.500)	(0.795)
N	552	373	416
r2	0.103	0.283	0.274

* p<0.1, ** p<0.05, *** p<0.01. The base category for home language is: Igbo, Hausa, and other less frequently spoken languages. The base category for Teacher qualifications is no education or only primary and secondary education completed.

Table 9 OLS Regression of numeracy score factors

	(1) Bridge	(2) Private	(3) Public
Pupil is male	-0.079	-0.125	0.106
	(0.067)	(0.107)	(0.083)
Age	0.226	0.244	0.220
	(0.299)	(0.492)	(0.364)
Age squared	-0.013	-0.005	-0.008
	(0.020)	(0.032)	(0.021)
Home language is English	-0.119	-0.032	0.080
	(0.103)	(0.106)	(0.191)
Home language is Yoruba	-0.119	-0.128	-0.192
	(0.124)	(0.099)	(0.162)
Continuous household wealth index	0.025	0.023	0.006
	(0.026)	(0.040)	(0.061)
Head teacher total years of management experience	-0.016*	0.051**	-0.027
	(0.009)	(0.020)	(0.029)
Head teacher has university degree	0.129	0.463***	-0.200
	(0.120)	(0.113)	(0.150)
School management average score 1-5 (avg of 15 questions)	0.261**	-0.132	0.004
	(0.114)	(0.238)	(0.174)
School fees according head teacher	0.016	-0.006	
	(0.021)	(0.005)	
Teacher has a teaching qualification (no bachelors)	0.063	0.293	-0.033
	(0.123)	(0.230)	(0.194)
Teacher has a bachelors' degree	0.120	0.161	0.302
	(0.140)	(0.139)	(0.199)
Teacher motivation index (index of 29 Qs)	0.004	-0.025	-0.009
	(0.027)	(0.029)	(0.032)
Numeracy lesson seemed planned	-0.112	0.365**	0.313**
	(0.148)	(0.171)	(0.137)
Percentage of classroom time spent on teaching activity	-0.006	0.006*	-0.009
	(0.006)	(0.003)	(0.009)
Constant	-1.28	-2.196	-1.100
	(1.507)	(1.903)	(2.056)
N	464	399	439
r2	0.084	0.177	0.154

* p<0.1, ** p<0.05, *** p<0.01. The base category for home language is: Igbo, Hausa, and other less frequently spoken languages. The base category for Teacher qualifications is no education or only primary and secondary education completed.

3.4.4 Plausible bounds on causal effects using coefficient stability methods

So far we have made clear that we cannot make causal inferences with this observational data, as we don't know how students were assigned to schools. This means that we cannot distinguish effects of recent schooling on test scores from student family inputs and prior schooling. Controlling for observable student characteristics gets us closer towards a causal interpretation of the difference between groups, but there are still unobserved differences between students which may be important, and conflated with school type. The Oster (2017) method does two things, first it establishes the degree of selection (σ) of students into schools on unobserved factors that would be necessary for the true causal effect of the school type to be zero, in relation to the observed amount of selection on observable variables. Second, given an assumption that the degree of selection on unobservable factors is at most equal to the degree of selection on observed factors ($\sigma = 1$), it allows the presentation of a lower bound on the true causal effect. In our case, we are comparing the amount of selection due to pupil gender, age, family wealth, and home language, to other unobserved factors. The method is motivated by the idea that control variables have been purposively selected to explain as much of the variation in test scores as possible, such that they can be assumed to explain at least half of the variation for which it might be possible to explain. This approach has been used in estimating the effects of private schools using observational data by Sakellariou (2017).

We first estimate a simple regression of test scores on the school type indicator (in our case a dummy variable for Bridge), followed by the same regression with a full set of student controls (age, sex, a wealth index, and home language). We then observe the stability of the coefficient on the Bridge dummy variable (β) and the r-squared of each regression. Finally, we make an assumption about what would be a reasonable r-squared to expect if we had full information on the unobserved factors. Oster (2017) suggests a maximum potential R-squared value of 1.3 times the r-squared achieved in the full controlled regression.

We first present a set of comparison of Bridge with private schools and public schools, with and without controls, in the table below. The gap with private schools in literacy falls from 0.35 to 0.23 standard deviations after controlling for covariates. The gap with private schools in numeracy is small, negative, and statistically insignificant. The gap with government schools is large for both literacy and numeracy, with and without controls. Thus our main concern with coefficient stability is for the comparison with private schools for literacy, where controlling for student characteristics moves the coefficient towards zero. How much additional selection on unobservables would be required for the causal coefficient to be zero?

Table 10 Comparison of Bridge with Private and Public Schools

	Compared to Private Schools				Compared to Public Schools			
	Literacy		Numeracy		Literacy		Numeracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bridge	0.351***	0.233**	-0.055	-0.111	1.384***	1.375***	0.860***	0.911***
	(0.104)	(0.097)	(0.096)	(0.090)	(0.083)	(0.099)	(0.094)	(0.159)
Pupil is male		-0.163***		-0.132**		-0.131***		0.050
		(0.056)		(0.054)		(0.045)		(0.048)
Age		0.438**		0.484*		0.058		0.155
		(0.198)		(0.269)		(0.131)		(0.179)
Age squared		0.034***		-0.027		-0.006		-0.006
		(0.013)		(0.018)		(0.008)		(0.011)
Wealth quintile		0.071***		0.041**		0.039		0.024
		(0.024)		(0.022)		(0.020)		(0.034)
English spoken at home		0.088		0.047		-0.036		0.056
		(0.090)		(0.084)		(0.074)		(0.099)
Yoruba spoken at home		-0.123		-0.082		-0.173**		-0.016
		(0.084)		(0.085)		(0.072)		(0.100)
LGA Fixed Effects		Yes		Yes		Yes		Yes
Students	1,056	1,053	1,057	1,054	1,106	1,096	1,105	1,097
Schools	86	86	86	86	75	75	75	75
r2	0.046	0.128	0.001	0.053	0.441	0.499	0.190	0.231

* p<0.1, ** p<0.05, *** p<0.01 The base category for home language is: Igbo, Hausa, and other less frequently spoken languages. The base category for Teacher qualifications is no education or only primary and secondary education completed.

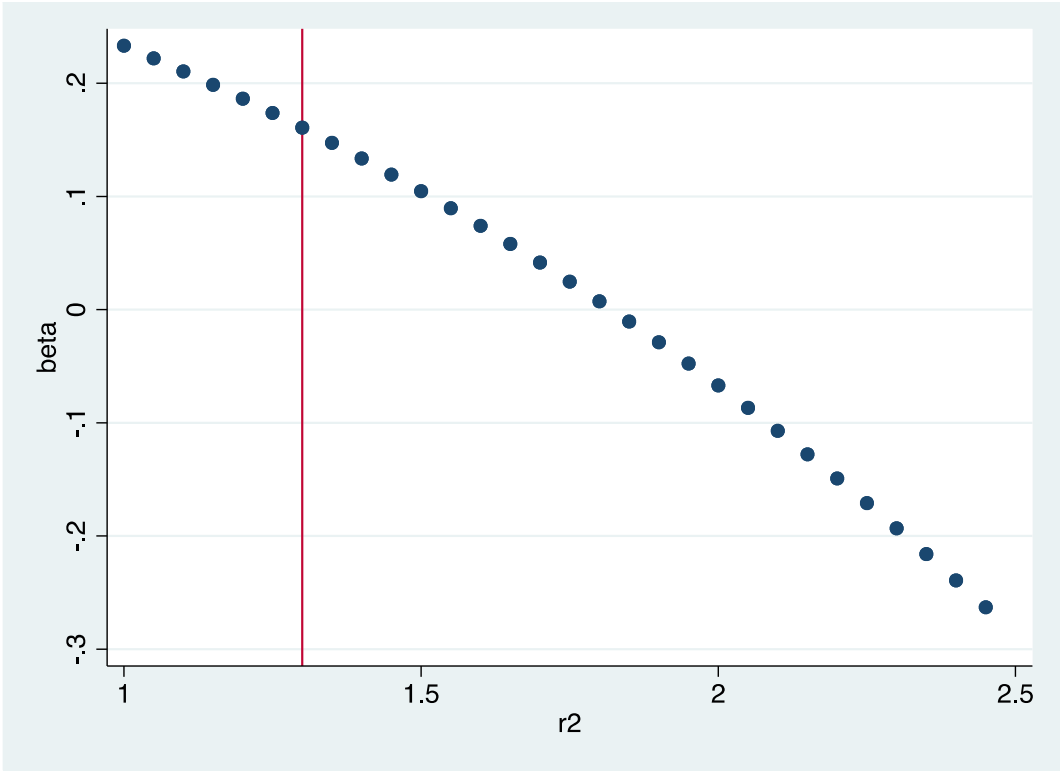
The table below presents estimates for the coefficient first assuming that the degree of selection on unobservables is equal (and in the same direction) as the degree of selection on observables $\sigma = 1$. Altonji, Elder and Taber (2005) and Oster (2016) suggest that a value of $\sigma = 1$ is a reasonable upper bound to expect. Given that control variables are typically chosen explicitly to try and control for variation, it may be reasonable to assume that selection on unobservables is no greater than selection on observables. In the case of $\sigma = 1$ the estimate for the coefficient is 0.161 standard deviations. We can also calculate the degree of selection on unobservables necessary for the true coefficient to be zero (delta, σ). This is 2.2 – that is, the selection on unobservables would need to be more than twice as large as the selection on observables for the true coefficient to be zero.

Table 11 Comparison of Bridge with private schools

	Literacy		
	Coefficient on Bridge Dummy	R squared	Selection
Regression with no controls (Table 10, Col 1)	0.351	0.045	
Regression with controls (Table 10, Col 2)	0.233	0.128	
Effect size if selection on unobserved variables = selection on observed variable (Beta, β if σ = 1)	0.161		1
Required relative selection on unobservables (Delta, σ) for true effect size to be equal to zero (β = 0)			2.232
R Max (Assumed maximum possible R squared is equal 1.3 times the R-squared achieved in the regression with controls (Table 10, Col 2)		0.166	

This estimate of σ is based on an assumption that the maximum possible R-Squared is 1.3 times the R-squared reached in the fully controlled model. This value is proposed by Oster (2016) based on an analysis of 65 randomized studies from top five economics journals. 90 percent of the randomized studies are robust to a R Max of 1.3 times the r-squared reached in the fully controlled model. We demonstrate the sensitivity of the main estimate to varying the maximum possible r-squared in the figure below. For the true coefficient to be zero, the maximum possible R-squared would need to be 1.7 times the R-squared achieved in the fully controlled model.

Figure 35 Sensitivity of coefficient estimate to varying maximum possible r-squared



All of this suggests that the observed gap between Bridge and private schools in literacy is unlikely to be driven by selection on unobserved variables[30]. We do not apply the Oster method for public schools. This is because the gap is so large and remains large even with factoring in for observables, that it's highly unlikely that selection is driving the gap.

[30] Note that crucially, this only applies to which are related to the observables and share the same covariance properties as the observables. DFID'

4 Value for money?

An important question for parents and policymakers alike relates to whether additional money spent at Bridge or other private schools translates into additional learning. When thinking about the 'value for money' of a specific intervention or strategy, the objective is not to achieve the lowest costs but to maximise the impact of each dollar spent to achieve the desired 'quality' at the lowest possible price (Department for International Development, 2011). This question is best answered through a rigorous cost-effectiveness analysis (CEA). A CEA analysis requires a detailed analysis of school business models, costs of provision, and income from fees to arrive at the cost per pupil. In addition, a causal estimate of the impact of school type on learning is required (e.g. through an experimental or quasi-experimental comparison of learning outcomes). The detailed unit cost analysis and the causal estimate of learning are then combined to compute, for example, how much extra learning per unit of currency is delivered by different school operators.

For this paper, detailed data on each school's income and expenditure was not gathered, and the observational nature of the data does not allow us to clearly causally attribute learning achievement to different school types. We implement an alternative strategy to help us arrive at an indicative measure of cost per unit of learning: calculating the conditional mean literacy and numeracy scores per fees spent by parents. This method is adapted from Kingdon (1996) and Kingdon (2008), where predicted test score values were used in combination with provider unit cost to compute costs per test score point. Specifically, we ask: holding pupil background characteristics constant, how much learning is correlated with attending private schools versus Bridge per 5000 Naira[31] spent by parents? While there are ongoing per-pupil costs of education provision for all school types, we restrict the comparison to private schools and Bridge schools as the public school model is not based on charging fees, and calculating the cost per child at the provider level is beyond the scope of this analysis.

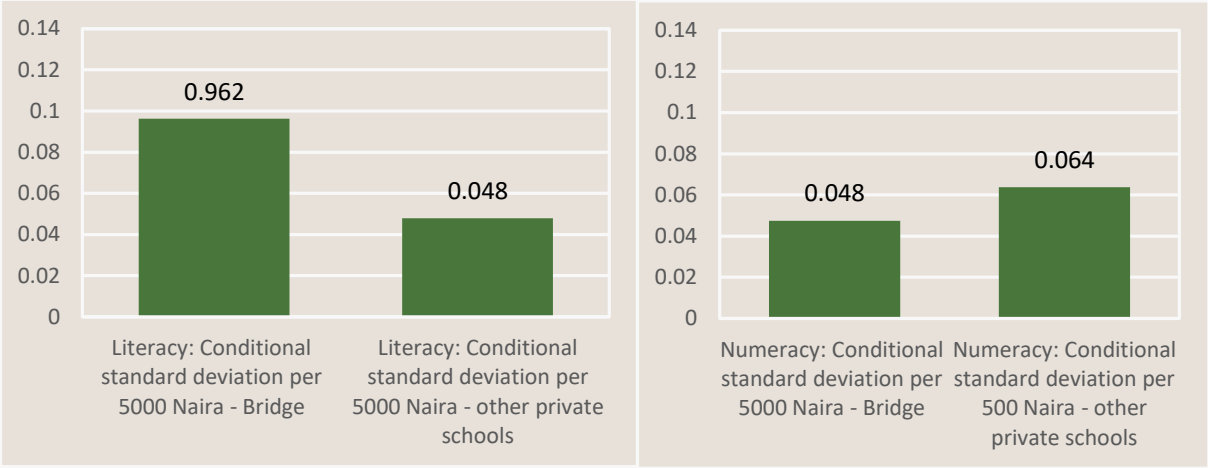
In interpreting these findings, it is important to bear in mind that only considering learning per fee spent by parent is different from considering learning per *full cost per child* at the school level. So, value for money from a parents' perspective is not the same as value for money from a policymakers' perspective. Therefore while we can provide an estimate of value for money from a parents' perspective, we cannot determine in this analysis whether it is the most cost-effective investment for a policymaker. Full cost per child at the school level and the cost-effectiveness of investing in comparable, feasible and context-appropriate alternatives are critical considerations.

We predict the scores for Bridge and other private schools from regressions of normalised literacy and numeracy scores on school type and student background characteristics which may determine selection into a specific school type - gender, age, household wealth, home language, and the local government authority in which a student resides. The regressions used were identical to those used in the Oster coefficient stability method discussed in the previous section. We then divide the predicted values by average fees at Bridge and other private schools and multiply by 5000 Naira, to arrive at the predicted standard deviation per 5000 Naira.

The figure below compares the conditional mean standard deviation per 5000 Naira in literacy and numeracy between Bridge and other private schools. In other words, it asks – holding student background characteristics which may determine selection into Bridge versus other private schools, how much learning does an additional 5000 Naira buy parents at Bridge and other private schools?

[31] We choose 5000 Naira because it is approximately 15% of the average fees paid, and because it is equivalent to approximately £10

Figure 36 Conditional standard deviation per 5000 Naira in Bridge vs other private schools



The figures suggest that 5000 Naira "buy more literacy" at Bridge compared to other school types (consistent with findings in Section 3.4.4), and less numeracy. Whilst we do not compute the standard error of the conditional standard deviation per 5000 Naira, findings from the previous sections suggest that the difference in literacy is likely to be statistically significant, whereas the difference in numeracy is not.

5 Conclusions

5.1 A summary of key results

1. Literacy, numeracy, and value for money

- In literacy, students at Bridge schools have better performance than students at other private schools (by 0.35 standard deviations) and public schools (by 1.38 standard deviations);
- In mathematics, students at Bridge schools have similar performance to students at other private schools and better performance than students at public schools (by 0.86 standard deviations);
- Adjustments for pupil age and gender, household wealth and home language, and local government area reduce differences, particularly in private schools - by 0.12 in literacy compared to private schools, by 0.01 in literacy compared to public schools, and by 0.03 in numeracy compared to public schools. However, they remain sizeable.
- Our findings suggest that Bridge may provide parents with better value for money in literacy. This does not necessarily equate with better value for money from a policymakers' perspective – determining whether investments provide taxpayers with value for money requires an impact estimate and more information on the cost and financial modelling of both public and private provision.

2. Students, teaching and management

- Students in Bridge schools and other private schools are less likely than students in public schools to belong to the bottom socioeconomic quintiles, to speak Yoruba at home, and to be overage.
- All school types enrol equal proportions of boys and girls.
- Teachers in Bridge schools report higher motivation than teachers in other schools, and Bridge schools are better managed than other schools. However, observed teaching practice does not differ substantially between school types for literacy. For numeracy, teachers in Bridge and public schools spend more time teaching than teachers in private schools.

3. Correlations with learning

- For literacy, students from better socioeconomic backgrounds have higher learning achievement in private and public schools, but not at Bridge schools. Numeracy achievement is not significantly correlated with wealth across school types.
- Girls perform significantly better than boys overall, and particularly in Bridge and public schools. The difference is fairly large – 0.16 standard deviations overall. In private schools, girls and boys perform similarly at the P2 level, in line with previous evidence from Lagos (EDOREN, 2016). We also find a statistically significant disadvantage for boys in literacy in public schools at the P4 level (EDOREN, 2017).
- Better school management is significantly correlated with higher literacy and numeracy achievement, particularly in Bridge schools and public schools.

- Teaching qualifications and bachelors' degrees are not significantly correlated with literacy and numeracy achievement.

5.2 Discussion and recommendations

The findings in this report highlight in particular the low levels of literacy and numeracy achievement in public schools at the primary 2 level, and the need to support these schools to increase their performance in both subjects. In addition, more research would help unpack why the stark differences in numeracy achievement in private schools compared to public schools found in this report are not as pronounced at the P4 level (EDOREN, 2017).

There are some positive indications for DFID's investment in Bridge: a Bridge education is correlated with better literacy achievement compared to other private school alternatives, beyond what we would reasonably expect to be the result of selection on observable and unobservable characteristics. However, we do not find significant differences in numeracy achievement between Bridge schools and other private providers in Lagos: a Bridge education is as good as that received in other low-fee private schools for numeracy.

Our findings suggest that Bridge may provide parents with better value for money in literacy, but not numeracy. This does not necessarily equate with better value for money from a policymakers' perspective, which requires an impact estimate and more information on providers' costs and financial models.

School management is strongly correlated with higher literacy and numeracy achievement. This suggests that programmes which seek to improve school quality assurance and inspection systems could include mechanisms to score schools on the 5 dimensions of good management captured in the tool: target-setting, operations, monitoring, people management, and leadership.

DFID should also seek to further understand the reasons why girls in Lagos have higher literacy rates compared to boys. In our ESSPIN study, we find that Lagos is the state with the highest literacy rates in Nigeria, and we find repeatedly that in Lagos girls have better literacy results than boys. There is likely to be much to investigate and learn from this example of positive deviance.

References

Allen, R. (2018). *CACSA report Bridge 1: Bridge pilot tests - numeracy.*

Allen, R. (2018). *CACSA report Bridge 4.1: Bridge tests - literacy v.1.1.*

Ashley, L. D. (2014). *The role and impact of private schools in developing countries. Rigorous Literature Review.* ASER Pakistan.

Aslam, M., Rawal, S., Kingdon, G., Moon, B., Banerji, R., Das, S., . . . Sharman, S. (2016). *Reforms to Increase Teacher Effectiveness in Developing Countries.* London: London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Bloom, N., Lemos, R., Sadun, R., & and Van Reenen, J. (2015). Does Management Matter in Schools? *The Economic Journal*.

Burgess, S., Davies, N., & Slater, H. (2009). *Do teachers matter? Measuring teacher effectiveness in England.*

Crawfurd, L. (2017). *School Management and Public-Private Partnerships in Uganda.* Research Into Systems of Education (RISE).

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*, 175-187.

Department for International Development. (2011). *DFID's Approach to Value for Money .* Department for International Development.

EDOREN. (2015). *Note on the reliability of children as survey respondents.*

EDOREN. (2015). *Note on the reliability of children as survey respondents.*

EDOREN. (2015b). *Developing Effective Private Education Nigeria (DEEPEN) - Mixed Methods Evaluation Report.*

EDOREN. (2016). *Developing Effective Private Education Nigeria Baseline Report.* EDOREN.

EDOREN. (2017). *Comparing Learning Outcomes in Public and Low- and Medium- Fee Private Schools in Lagos.* EDOREN.

EDOREN. (2018). *Bridge Fieldwork Implementation Report.*

Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrolments in states of India . *Demography*, 115-132.

Glewwe, P., Hanushek, E., Humpage, S., & Ravina, R. (2013). *School Resources and Educational Outcomes in Developing Countries: A Review of the Evidence from 1990 to 2010.* NBER Working Paper Series.

Hanushek, E., & Woessmann, L. (2011). *Overview of the symposium for performance pay for teachers.*

Härmä, J. (2013). Access or quality? Why do families living in slums choose low-cost private schools in Lagos, Nigeria? *Oxford Review of Education*, 548-566.

Kingdon, G. (1996). The quality and efficiency of private and public education: a case-study of urban India. *Oxford Bulletin of Economics and Statistics*, 57-82.

Kingdon, G. S.-s. (Chakrabarti, Rajashri (Ed.); Peterson, Paul E.(Ed.)). School-sector effects on student achievement in India. In *School choice international: exploring public-private partnerships.* Cambridge: MIT Press.

Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer?. Review of Income and Wealth. *Review of Income and Wealth*, 128-165.

Krieg, J. M., & Storer, P. (2006). How much do students matter? Applying the Oaxaca decomposition to explain determinants of adequate yearly progress.". *Contemporary Economic Policy*, 563-581.

Lemos, R., & Scur, D. (2017). *Developing Management: An Expanded Evaluation Tool for Developing Countries .* Research Into Systems of Education.

Muralidharan, K. (2013). *Priorities for primary education policy in India's 12th five-year plan.* India Policy Forum (Vol. 9, No. 1, pp. 1-61). National Council of Applied Economic Research.

Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 1-18.

Oxford Policy Management. (2018). *A diagnostic of early learning in Liberia.* Oxford Policy Management.

Romero, M., Sandefur, J., & Sandholtz, W. A. (2017). *Outsourcing Education in a Fragile State: Experimental Evidence from Liberia.* https://sites.tufts.edu/neudc2017/files/2017/10/paper_484.pdf.

Sakellariou, C. (2017). Private or public school advantage? Evidence from 40 countries using PISA 2012-Mathematics. . *Applied Economics*, 2875-2892.

Singh, R., & Sarkar, S. (2012). *Teaching quality counts: how student outcomes relate to quality of teaching in private and public schools in India.* Oxford: Young Lives.

Tooley, J. &. (2013). *School choice in Lagos state. .* DFID.

Tooley, J., & Yngstrom, I. (2013). *School choice in Lagos state. .* DFID.

Tooley, J., & Yngstrom, I. I. (2014). *School Choice in Lagos State - Final Summary Report.*

Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health policy and planning*, 459-468.

Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008). Investigating ceiling effects in longitudinal data analysis. *Multivariate behavioral research*, 476-496.

World Bank. (2017). *The role of the private sector in Lagos, Nigeria.* World Bank.

Yngstrom, I. (2014). *DEEPEN Poverty Paper.* DEEPEN.

Annex A Learning scale score construction

A.1 Background

This annex discusses the learning scale score construction for literacy and numeracy. Section A.2 clarifies the constructs measured. Section A.3 details the item targeting strategy. Section A.4 summarises the psychometric analysis conducted.

A.2 Clarifying constructs

The first step in undertaking a learning measurement exercise is to define exactly what is to be measured. Defining exactly what it is that we expect children to know and be able to do is at the heart of the measurement of learning. Literacy and reading are quite different constructs, as are numeracy and mathematics. At its simplest, to be literate means to be able to observe symbols or signs from a page and to ascertain meanings that are standardised to some extent. It also means being able to produce text by writing the same symbols or signs in order to record meaning. So, even in its most basic of forms, literacy is a lot broader than reading. International gold-standard assessment programmes, such as Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS), go even further, incorporating an element of social context into their understandings of literacy. Similarly, the OECD's definition of numeracy is broader than being able to operate mathematical equations – rather, it hinges on the *'ability to access, use, interpret, and communicate mathematical information and ideas, to engage in and manage a range of situations in life'* (OECD, 2014). This definition is broader than understanding mathematical concepts but also includes being able to apply them in a variety of known and unknown situations.

The constructs measured as part of this comparison study were English literacy/language and numeracy. The study uses the broader constructs of literacy and numeracy, rather than reading and mathematics, as discussions with stakeholders during benchmarking workshops for EDOREN projects (i.e. ESSPIN, TDP, GEP3, and DEEPEN) have highlighted that the programme expects children to improve learning in areas that fit into these broader areas of literacy and numeracy – such as to understand texts, use reading to understand the world, draw inferences and communicate in writing, use money in everyday life, and read a clock.

A.3 Targeting

The flawed assumption that tested children are able to read or write already has been avoided. Test items were designed or selected from existing EDOREN tests to measure skill levels below, at and above the skills assumed to have been reached given the grade the child is attending.

The major weakness in data measuring literacy and numeracy in low and middle-income countries is that assessments measure skills at levels that are too high for most of the children taking the tests. This was observed in the case of Mali, where over 80% of students in Grade 2 could not read a single word in four national languages, while in the Nigerian state of Sokoto 81% of the students could not read full words (Global Education Monitoring, 2012). These floor effects are not problematic if the goal of the assessment is to establish national learning levels. However, one of the goals of this study is to act as a baseline for an evaluation. In the presence of severe floor or ceiling effects, it would not be possible to detect growth in literacy or numeracy unless those children previously out of range of the tests come into the ability level measured by the test. For those children who do not come into range, but whose literacy or numeracy levels do improve, the impact of the intervention will be underestimated. Therefore, the assessments were

adapted to ensure item difficulty matched pupils' ability. The level of the assessment and pupils' abilities were assessed during the piloting and adjusted based on the results.

In the case of Lagos, ceiling effects were observed during the DEEPEN baseline evaluation and therefore, more difficult items needed to be added to the instruments in order to ensure the test was targeting to the higher proficiency levels of the pupils in Lagos. This was done using IRT. The tests were placed onto the same scale using link items to link the difficulty of the ESSPIN test to a new version of the test and to place the results onto the same scale (metric).

A.4 Psychometric analysis

The use of Rasch modelling (IRT) increases the amount of analysis required as more sophisticated techniques are used, rather than adding up a total number of items correct in the test and converting the number into a percentage score. In this regard the evaluation team followed the following steps: the first step was to test the psychometric properties of the items to ensure they were useful measures of what pupils know. The second step was to remove any items that did not perform well. In a third step the team ranked the items according to difficulty. This was done by a psychometrician, using fit-for-purpose software. The software then also ranked pupils according to their ability and placed the pupils and the items onto the same metric. This is a probability model as pupils are placed on the scale according to the probability of a pupil answering the corresponding item correctly.

A non-parametric IRT analysis was used to see the extent to which items met basic assumptions required by the usual parametrised IRT analysis. For both literacy and numeracy, the items were considered satisfactory. Both the literacy and the numeracy data was analysed using one-parameter IRT model. The one parameter model has an Expected A Priori (EAP) reliability of 0.82 for literacy, and 0.88 for numeracy. These are acceptable values (Allen 2018a, Allen 2018b).

Full discussions of the psychometric analysis for literacy and numeracy, in the pilot and in the survey, are available in separate technical reports.

Annex B Adapting the World Management Survey Tool

We measured school management quality using an adapted version of the Bloom et al. (2015) tool adapted for Uganda by Crawfurd (2017). We carried out interviews with head teachers and administered 15 questions covering five main areas: target-setting, monitoring, operations (planning and leading teaching), people (teacher) management, and leadership. These are detailed below.

- 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 ongoing 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;

Leadership: how the school's vision is set.

Adaptation process

- We piloted the tool adapted for Uganda in January 2018, to check head teachers' understanding and enumerators' skill in administering the questions;
- We followed Lemos & Scur (2017) and expanded the survey "horizontally" (by introducing half scores) to allow for greater variation of scores and finer differentiation between levels of management.
- We revised the scoring rubric in line with Lemos & Scur highlighting that:
    - A higher score must be awarded the more structured and formalised the process in place (e.g. a school where teachers being hired on an ad-hoc basis would receive a score of 2, whereas a school where there is a formal interview process would receive a score of 3);
    - A higher score must be awarded the more proactive the approach to management (e.g. a reward system for teachers is in place versus a reward system is in place and active consideration is given to ensuring that this system is fair).
- We included indications of frequency (e.g. once a year for regularity of measuring student results for a score of 3, and once a term for a score of 4) to reduce the subjectivity of enumerator rankings.

Each interview was double scored: the first interviewer was accompanied by a second interviewer whose main role was to monitoring the quality of the interview being conducted by taking notes and separately scoring the responses after the interviews had ended. The first and second interviewers would then discuss their individual scores to correct for any misinterpretation of responses. All interviewers held a debrief at the end of each day where each group explained their scoring for each question. Any differences in opinion were voiced and scores were calibrated with the help of the survey manager who had received in-depth

training into the tool. We mixed pairs of interviewers as much as possible throughout the survey, conditional on geographic limitations.

Annex C Household wealth index

C.1 Background of the analysis

This study aims to determine the differences in learning levels of P2 students in Bridge, other low- and medium-fee school pupils, and public school pupils in Lagos. In line with previous studies, our basic premise is that an assessment of learning outcomes necessitates appropriately controlling for the socioeconomic status of the household of each child. It is also relevant to understand the overall level of welfare of households whose children attend private and public schools. One of the approaches that could be used to assess the poverty level of these children's households accurately could entail conducting a household visit and administering a detailed survey enquiring about household food and non-food expenditures. These surveys would, however, represent a logistical problem for this study: their complexity rules out the possibility of administering them to children and hiring enumerators to visit all households would be too expensive and time-consuming. The alternative option of administering the asset questions to children's parents carried the risk of not having any background information for the parents that could not be found and interviewed for the survey. Therefore the team opted to obtain a measure of socioeconomic background from a relatively simple set of questions that a child can answer and that is adapted to the Lagos-specific context to ensure that it captures enough variability in its measure. This method has been successfully validated in two other EDOREN studies: DEEPEN and the P4 Comparison Study, and a technical write-up demonstrating a high level of consistency between children's and parents' responses from the DEEPEN study is available.

C.2 Index Methodology

We created a socio-economic status index based on coefficients derived from a polychoric principal component analysis (PCA) on our data. We use polychoric PCA instead of the simple PCA approach popularized by Filmer and Pritchett (2001) in order to correct for the fact that simple PCA with a collection of ordinal variables might produce spurious correlations, as demonstrated in Kolenikov (2009). However, the essence of the technique is the same: PCA looks for the linear combinations of variables that capture the information they jointly convey most successfully.

This method of estimating wealth has the advantage of capturing longer-term socio-economic status that regular consumption and expenditure surveys. There is a strong rationale that the assets included in a household wealth index analysis are more relevant in education research as they are more likely to capture household investment over time. The method is not, however, without flaws. Re-test validity can be low, and PCA is very sensitive to the choice of variables included in the analysis (Onwujekwe et al., 2006). With these caveats in mind, the alternative – a consumption survey – is also often inaccurate, prone to recall error and seasonal effects, and substantially less cost-effective (Vyas and Kumaranayake, 2006).

C.3 Assets included in the analysis

In implementing the survey, a separate list of twenty questions related to household wealth was administered to children. These have been mostly yes/no questions about simple assets such as a TV and a fridge, as well as simply-phrased questions about parents' education[32] and religion. The questions were based on a combination of our work constructing wealth indices for the DEEPEN baseline survey and the P4 Comparison Study (EDOREN, Developing Effective Private Education Nigeria Baseline Report, 2016;

[32] We have excluded the education of the household head from inclusion in the index because the variable was highly correlated with the final wealth index, and therefore could not have been used as a separate predictor in our regression analysis.

EDOREN, 2017). We have excluded questions which were difficult for the children to answer in the DEEPEN baseline survey, such as type of roof material in children's house, or source of drinking water. We have also excluded questions which in our previous experience had very low explanatory power (such as number of mosquito nets). We have, however, expanded the set of questions to include simple yes/no questions about assets from the Education Sector Support Programme in Nigeria (ESSPIN), which had previously successfully captured wealth levels among public school children. This means that our index is based on a comprehensive list of assets obtained from our work with both private and public schools. The assets included in the index are listed in the table below.

Table 12 Assets included in the construction of the index

Dimension	Level
Religion	Christian
	Muslim
	Other
Toilet	Own flush
	Shared flush
	Other
Room per person (excl. bathroom, kitchen)	$X \leq 0.21$
	$0.21 < X \leq 0.35$
	$0.35 < X \leq 0.5$
	$0.5 < X$
Fridge	None
	One or more
Car	None
	One or more
TV	None
	One or more
Radio	One or more
	One or more
Sofa	None
	One or more
Chair	None
	One or more
Table	None
	One or more
Bed	None
	One or more
Mat	None
	One or more
Sewing machine	None
	One or more
Bicycle	None
	One or more
Generator	None
	One or more

Fan	None
	One or more
Computer	None
	One or more
Mobile phone	None
	One or more
Air conditioner	None
	One or more

Annex D Sampling

D.1 Background

The aim of the sampling strategy adopted was to enhance the comparability of the different groups of schools included in the study: Bridge, private and public schools. In order to achieve this, we matched Bridge schools with low- and medium-fee private schools derived from EDOREN's recent P4 study, which compared private and public school P4 pupils in Lagos. A location-specific random sub-set of public schools was then drawn from the P4 sample to create the public school arm for this study.

D.2 Private school matching

The first step of the matching exercise was to clean and merge two sets of P4 data: the original school sample and the school replacement sample. From the complete private schools and Bridge school data, we then created a composite score index based on four variables: nominal school fees, total number of pupils and GPS coordinates (i.e. longitude and latitude). Specifically, the following steps were followed:

- A standardised (Z) score was calculated for each of the four variables;
- The four individual Z scores were summed up into one variable and that variable was again standardised;
- The resulting variable was the composite index based on four equally weighted standardised components.

The Bridge school and Private school dataset now had individual composite variables, which were comparable given the standardisation protocol. Using the composite variable, a distance variable was calculated. The distance variable identified one or more private schools closest to a Bridge school based on the value of the composite score of the two sets of schools.

From the returned list of matches, we selected the private school with the smallest distance to a Bridge school, dropping the other matches for the same schools. This automatically retained 33 Bridge schools matched to 33 private schools based on the smallest calculated distance. We removed the 33 matches from the full sample and manually selected 4 schools that were not automatically matched in the data. This is because the automatic matching found multiple private schools being closest to the same Bridge schools. However, we wanted and needed (for power considerations) to match all Bridge schools to at least one private school. The manual selection was done by checking for the private school whose composite index score was closest to the 4 Bridge schools that were not automatically matched. The sample for the private school arm of the study thus includes 37 private schools matched to the 37 Bridge schools.

D.3 Public school sampling

Public schools were drawn randomly for the P4 study from the Lagos sample of ESPINN schools. The sampling strategy adopted for both private and public schools in the P4 study was made as consistent as possible (e.g. using implicit stratification by location and a random selection approach proportional to school size) in order to enhance comparability. This allowed us in turn to select a random sub-sample of P4 public schools for this study. Finally, to further improve comparability and given the importance of location considerations, the public school sampling frame for this study was restricted to the LGAs in which the private schools matched with Bridge schools are located.

EDOREN

Education Data, Research and Evaluation in Nigeria

No. 2, 16 Mafemi Crescent,
Utako, Abuja, Nigeria
Phone: +243 (0)810 727 8718 | +234 (0)817 667 8243
Website: www.edoren.org | Email: info@edoren.org