The EdoBEST Effect

Joan Oviawe, Steve Cantrell, Lisa Chen, Mark Buttweiler, Spencer Nash, Stacey Nwokeyi

With a Foreword by Governor Godwin Obaseki

STATE UNIVERSAL BASIC · EDO STATE · SUBEB · EDUCATION BOARD · QUALITY EDUCATION FOR ALL

EdoBEST · BASIC EDUCATION SECTOR TRANSFORMATION

An Edo SUBEB Initiative

The EdoBEST Effect

Joan Oviawe, Steve Cantrell, Lisa Chen, Mark Buttweiler, Spencer Nash, Stacey Nwokeyi

With a Foreword by Governor Godwin Obaseki

February 2019

Joan Oviawe, Chairman, State Universal Basic Education Board, Edo State, Nigeria.
Steve Cantrell, Vice President, Measurement & Evaluation, Bridge.
Lisa Chen, Director, Measurement & Evaluation, Bridge.
Mark Buttweiler, Manager, Measurement & Evaluation, Bridge.
Spencer Nash, Analyst, Measurement & Evaluation, Bridge.
Stacey Nwokeyi, Project Director, EdoBEST.

Foreword
by Governor Obaseki
Edo State

It is generally accepted that education provides the golden ticket to a better life. However basic education which is the education offered at the primary school level is the foundation and corner stone of any education system. It provides children with the opportunity to learn, explore and understand new perspectives. Quality basic education puts children on the part to a secure future and the satisfaction of contributing to both their own economic well being and that of their families and communities as well

This is why basic education is one of the most critical priority areas for my administration. The aim of our education policy is to reverse the decline in education quality and standards in Nigeria over the past decades.

Our commitment to restore education in Edo State back to its former glory, led us to launch the **Edo Basic Education Sector Transformation (EdoBEST)** program in April 2018. EdoBEST leverages the transformative power of cutting edge technology to deliver outstanding learning outcomes to primary school children across Edo State by better supporting their teachers. The program started with a pilot where 1,500 EdoBEST teachers were trained and equipped with the tools and skills to execute a unique type of learning experience for children which though unusual, but is however much needed in public education in Nigeria. The impact on the children of Edo State will no doubt be exponential and we hope continue to see the multiplier effect far into the future.

I am very pleased that there is already evidence of the impact which this innovation is making in the lives of our children, as we monitor and evaluate the results from EdoBEST.

This study, which is aptly titled **The EdoBEST Effect**, details the tremendous strides that have already been made which include equipping teachers in participating schools with computers, providing mobile devices for administrators and the use of technology to strengthen teaching practice and improvements in operational management of the schools. The project has achieved significant gains in learning for remedial content.

However, we cannot rest on our laurels as this is only a first start and much still needs to be done. The report also highlights areas where improvement is still required such as the need for an examination of our curriculum and school schedule given that pupils are 2-3 years behind the expected grade level. How do we maintain our standards while giving more focus to building core literacy and numeracy skills that are the foundation for learning more advanced concepts? The answers requires both an inward look at our policies and continuous investment in education, and an increased focus on creating an enabling environment that attracts the best teacher talent to Edo State.

The rigor that has been applied to conceiving, executing and monitoring EdoBEST is commendable. It represents best-in-class public private partnership, and the standards of accountability that ensures excellence.

For this spirit of excellence, I would like to commend and thank Dr. Joan Osa Oviawe, the Chairman, Edo State Universal Basic Education Board (SUBEB) and her entire team. They have shown a great resolve to effect all that is required to usher in the most radical changes to Edo education in a generation. My gratitude also goes to Bridge International Academies, our technical partners, who have availed Edo State of their renowned ICT-enabled methods. They have devoted their time and resources and demonstrated heartwarming faith in our vision.

My final thanks go to all the teachers, children and parents of the EdoBEST schools, for their trust in us. I also commend their willingness to work with the Government and our technical partner, quickly putting aside any ambivalence as we change from status quo.

In conclusion this report serves as an initial appraisal of our watershed decision to change the educational trajectory of our state and nation. We realise that the early start progress we have made needs to be built upon and sustained as we learn lessons from what we have achieved and look with confidence into the future.

Mr. Godwin N. Obaseki
Governor, Edo State

Executive Summary

The Edo State Universal Basic Education Board (SUBEB) launched the Edo Basic Education Sector Transformation (EdoBEST) initiative in April of 2018. 263 schools enrolling over 31,500 pupils participated in the pilot. This study of 30 EdoBEST pilot schools revealed several encouraging findings: Pupils learned more, spent more time learning, worked harder, and experienced a more positive classroom environment. Girls in EdoBEST schools outperformed all other students. These early results, based upon a one-term pilot, are cause for optimism as the program continues into next school year and expands to more schools and pupils.

Sample

Our pilot study examined pupil performance and a range of teaching and learning indicators in 30 EdoBEST schools and 30 schools not yet in the EdoBEST program, which we label status quo schools. These 30 EdoBEST and 30 status quo schools are representative of Edo State schools. They were selected to be characteristically similar to Edo State schools and each other. We examined Primary 3 (P3) classrooms, the highest grade participating in the EdoBEST pilot. Because these schools also teach Primary 4 (P4) pupils, none of whom received EdoBEST instruction, we were able to use results from P4 classrooms to estimate the differences that would have existed, in the absence of the EdoBEST program, between the EdoBEST participating schools and the status quo schools.

Method

We used a difference-in-difference method to attenuate selection bias. In our study, any preexisting differences between EdoBEST and status quo schools could threaten the validity of our findings. An uncontrolled study would only be able to observe differences between the EdoBEST schools and the status quo schools. In an uncontrolled study, we would not be sure how much of the observed differences to attribute to the EdoBEST program and how much to attribute to pre-existing or non-program related differences. Using the difference-in-differences method, we do not need to establish that the groups of schools are identical. Instead, we estimate any preexisting differences and carefully adjust for them. This adjustment provides more accurate impact estimates for the EdoBEST program.

(A) P4: Existing, non program-related differences

(B) P3: Pupil performance differences

(C) Adjustment: (B)-(A)

EdoBEST Effect
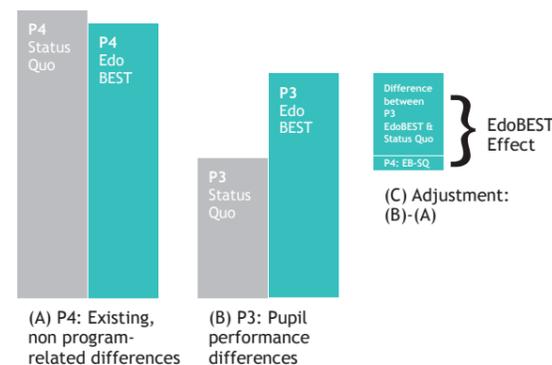
Our study uses the P4 classes as a control. Since the program was not taught in P4 classes, we can compare P4 classes in EdoBEST schools with P4 classes in status quo schools. Any differences in the P4 classes we observe are *not* caused by the program and, therefore, are useful approximations of existing, non-program related differences between the P3 classes. When we subtract non-program related differences (differences between P4 classes) from the total differences (differences between P3 classes), we are left with a good estimate of the EdoBEST impact.

The Big Finding

The big news is pupils have performed better in EdoBEST classrooms, even at this early stage of implementation.

There is still much work to do. The assessment results suggest most pupils, across all schools, are unable to perform grade-level tasks in literacy and numeracy. The average score on both the literacy and numeracy assessments barely exceeded the score a pupil would have received by guessing the answer to each question. Moreover, we observed a narrow distribution of scores around this average. This pattern suggests most pupils struggle with grade-level content. This finding has implications for the state's curriculum, which may need to adopt a "teaching at the right level" strategy to fill gaps in pupil understanding alongside introducing grade-level content.

Even so, the EdoBEST program is working. EdoBEST pupils significantly outperformed status quo pupils on remedial ("revision") items. Since we administered the same assessment to P3 and P4 pupils, we use the performance differences between P3 and P4 pupils to estimate one year of learning. In the status quo schools, the P4 pupils answered correctly 10% and 11% more items in revision math and literacy, respectively, than P3 pupils. This becomes our rough estimate for a year of learning. In the EdoBEST schools, the P4 pupils (none of whom received EdoBEST instruction) answered only 3% more items in maths and 5% more items in literacy than P3 pupils (all of whom received EdoBEST instruction). The difference in gains achieved by the P3 EdoBEST pupils is therefore 7% in maths and 6% in literacy. This equates to nearly three-quarters of a year more maths instruction and nearly two-thirds of a year more literacy instruction. Furthermore, these impacts are driven by girl pupils.

While it is tempting to trumpet these impacts as enormous, we are quite aware of the vast gap existing between pupils' actual performance and the expectations set by the national curriculum. Even so, EdoBEST does appear to be an accelerator. If the single term gains are multiplied over the full three-term year, EdoBEST appears to be a promising approach to raising pupil performance to the level of the national curriculum.

	Math					English Literacy				
	EdoBEST Effect		P3 vs. P4 Gap		EdoBEST Effect as % of Gap	EdoBEST Effect		P3 vs. P4 Gap		EdoBEST Effect as % of Gap
	In % Correct	In SDs	In % Correct	In SDs		In % Correct	In SDs	In % Correct	In SDs	
Revision	7%*	0.27*	10%	0.38	71%	6%*	0.33*	11%	0.61	54%
At-Level	2%	0.12	5%	0.35	34%	2%	0.08	10%	0.45	18%
Total	4%+	0.24+	7%	0.43	56%	5%	0.25	11%	0.61	41%

Note: The P3 vs P4 gap is calculated as the mean difference in standardized scores between P3 and P4 pupils in status quo schools.

EdoBEST schools provide more …

Training ↑ 73%

Teacher Tech ↑ 100%

Textbooks ↑ 9 books per class

Teachers more often …

Stay for the Full School Day ↑ 45 min.

Use Praise ↑ 54%

Praise More Pupils ↑ 34%

Pupils more regularly …

Follow Directions ↑ 25%

Focus ↑ 21%

Work Hard ↑ 45%

Have Work Checked ↑ 27%

This results in …

1 + 2 = 3

Improved Literacy Revision Scores ↑ 6%

Improved Maths Revision Scores ↑ 7%

which is equivalent to …

71% more schooling in mathematics revision

54% more schooling in English literacy revision

… or learning in 1-term what would otherwise take 1-year.

Theory of Action

We follow the EdoBEST program logic for measurements and analysis: Provide teachers with strong feedback, well-designed instructional guidance, national standards-aligned materials and texts, and more scheduled instructional time. These supports, in turn, strengthen teachers' abilities to create a more productive classroom and help pupils stay focused, follow directions, and work hard. These habits, in turn, help pupils learn better.

The figure below depicts our model and reports the single term benefits of EdoBEST. Teachers at EdoBEST schools received more training, used the provided technology, and had more textbooks in their classrooms. Their behavior also differed from teachers at status quo schools. EdoBEST teachers spent more time at school, provided more feedback to pupils, used praise and distributed it liberally to many pupils, rather than concentrating solely upon top performers. There is also evidence of positive spillovers on teachers who did not directly participate, but who were managed by headmasters who participated in the EdoBEST program: teachers indirectly exposed to EdoBEST were less likely to engage in corporal punishment. EdoBEST pupils responded differently than their peers at status quo schools. EdoBEST pupils more often followed directions, stayed better focused, and worked harder. All of this led to more learning.

Final Thoughts

The EdoBEST pilot was successful across multiple measures. Even so, we recognize that a single term is too brief to judge the success or failure of a program. What we did observe, even in this short period, was a significant gap between the curriculum expectations and the preparedness of pupils. While this study was limited to a single grade, it is reasonable to assume that the EdoBEST effect was not limited to P3 pupils, but extended to all younger pupils as well. If this pilot were a three-term pilot, rather than a single term, it is not hard to imagine that the P3 EdoBEST pupils would have surpassed the P4 pupils. When we extend that logic to earlier grades, we can be cautiously optimistic that EdoBEST is already working to close the large preparedness gap among the youngest pupils. This is merely speculation at the moment, but it is also our strong recommendation that, during the next phase of the program, the Edo SUBEB explore the preparedness gap and whether the EdoBEST instruction is at the right level for pupils at each grade.

Contents

I. Introduction

Over the past decade, economic growth has improved across Sub-Saharan Africa. Access to schooling has broadened; the gender gap in enrollment is closing. Nonetheless, significant challenges lie ahead in the education sector, the foundation of prosperity. Struggling behind other regions in Sustainable Development Goals, Sub-Saharan Africa still has the highest proportion of out-of-school children in the world. [1] Furthermore, nearly 1 in 3 pupils never complete primary school.

Nigeria is no exception. It has the greatest potential of becoming the economic powerhouse of Africa, but it is comprised of many states with some of the lowest primary school enrolment and completion rates in Sub-Saharan Africa. Nearly one-third of school-aged pupils never attend school.[2] Half of Nigerian children are not engaged in learning because they cannot read or write, and the vast majority in the lowest economic quartile are illiterate.[3] Such poor learning outcomes disincentivize attendance, particularly against the opportunity cost of earning much needed additional income.

A. State of Education in Edo

Edo is a state in southern Nigeria where 60% of the population live below the poverty line. Approximately 23% of the population are primary school-aged children in need of a better education.[4] Absenteeism is high and learning is low: 6.9 days (or 44%[5]) of school are missed each month; approximately 1 in 4 Edo children aged 5-16 are illiterate and 1 in 5 innumerate. The primary school completion rate is 45%, falling well below the national average of 68% for Nigeria, and 98% for OECD member countries.[6]

Figure 1. Primary School Completion Rate by State and World Region (2016)[7]

---
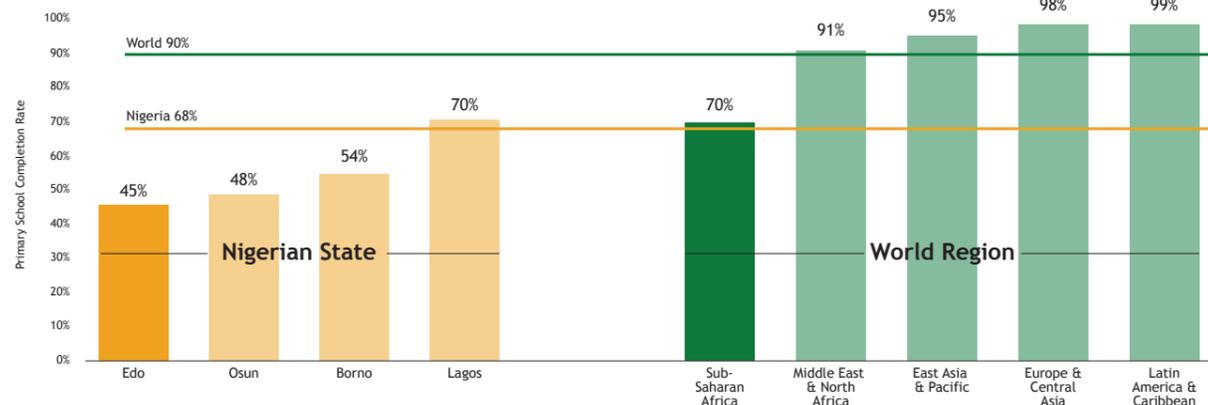
1 2015 MDG Report.

2 2010 Nigeria Education Survey.

3 "Nigerian children to benefit from Bridge International Academies," BusinessDay.

4 There are approximately 907,857 children ages 6-14 years old in Edo. 2016 Nigeria Education Indicators and Nigeria Data Portal (web).

5 Based on the 2018-19 EdoBEST academic calendar, there are 162 active days in the school year across 10.4 months, excluding holidays, breaks, exam marking days, and parent-teacher conferences. This results in an average of 15.6 days per month during the school year.

6 Nigeria Education Indicators 2016 and UNESCO Institute of Statistics.

7 Ibid. Primary school completion rate is calculated as the number of entrants to the last grade divided by the number of school-aged children for that grade.

B. Embarking on EdoBEST

The Edo State Universal Basic Education Board (SUBEB) launched the Edo Basic Education Sector Transformation (EdoBEST) initiative in April of 2018. EdoBEST aims to develop a highly skilled, supported, and motivated teacher workforce that, when coupled with improved infrastructure and integrated school management systems, will deliver learning and restore confidence in public schools.

The urgent need to build capacity for over 15,000 teachers and ensure learning for more than 320,000 pupils prompted a partnership with Bridge International Academies (Bridge). Bridge works with SUBEB to build local capacity, provide teaching and learning resources, and implement teacher development programming. The program pilot took place during Term 3 (April – July) of the 2017-18 school year for Primary 1-3 pupils.[8] The partnership is expected to increase both the number of participating schools and grades in 2019, reaching all ECCDE, Primary and Junior Secondary Schools in Edo State. During the four-year program, Bridge will transition its full field support responsibilities to SUBEB, supporting SUBEB in establishing Quality Assurance and Leadership & Development teams that are expected to continue using the EdoBEST data dashboards, make evidence-based decisions, and use effective coaching. Through extensive capacity building and knowledge sharing, the EdoBEST program seeks to ensure the transformation of the status quo across the state, with SUBEB firmly in charge and responsible for internalizing the process and practice changes integral to the EdoBEST program. After the transition of field support to SUBEB, Bridge will continue to support school leader and teacher training, provide extensive supplementary materials, and integrated technology solutions to ensure SUBEB has visibility on teacher and school performance.

---

8 Given time and resource limitations, the pilot focused on early primary grades. Schools were classified as EdoBEST if they had at least one headmaster and one teacher successfully complete training. Not all schools had participating teachers for each Primary 1, 2 or 3 classroom.

Specifically, Bridge provided the EdoBEST initiative with the following inputs:

- **Learning resources and materials** including textbooks, independent study books, individual homework books, lesson guides for teachers via a tablet, and a standardized daily timetable for each grade
- **Training and ongoing support** to better equip teachers for effective classroom management and quality delivery of the Nigerian national curriculum
- **Technology supported management and quality assurance systems** to help headmasters become more effective managers to attain greater learning outcomes for pupils
- **Building institutional capacity** and supporting SUBEB organizational structures to ensure EdoBEST sustainability following the end of the partnership

From 5th April to 19th April 2018, more than 1,500 teachers and headmasters from 263[9] schools across 18 LGEAs participated in the inaugural EdoBEST training. Teachers were trained in (1) the vision of a teacher, a pupil, and the EdoBEST program (2) classroom management practices to motivate and encourage participation by pupils, (3) use of the tablet as a teaching resource (lesson guides) and tool to track pupil attendance and learning, (4) techniques to deliver concrete and actionable feedback to all pupils, and (5) how to generate a positive, nurturing, and safe classroom environment.

Once teachers returned from training, an experienced teacher and leadership development officer visited each school every 1-2 weeks, providing coaching and individualized feedback. In addition, a quality assurance officer would visit every school approximately once every 2-3 weeks to ensure operations were running smoothly and help troubleshoot any issues.

The initiative also experienced challenges, including a teacher shortage, too few classrooms, poor network coverage and connectivity, and technology adoption barriers.

C. Learning Early to Understand What is Working

The EdoBEST initiative is not an isolated intervention focused on one area of education, but rather a whole-system reform with many integrated components. The difficulty of implementing system changes within a few short months calls for a comprehensive understanding of which components are working, to what extent, and how. Furthermore, it is critical for such learnings to occur quickly during the Phase 1 pilot, in order to inform implementation of future phases.

The theory of change underlying EdoBEST is summarized in Figure 2 below. Bridge directly provides resources, capacity building, and quality assurance processes (inputs)[10], which, combined with existing talent at schools, will increase learning time, improve the learning environment, and raise instructional quality (drivers). Together, these drivers expand pupils' opportunity to learn, and through consistent and persistent efforts, will result in learning gains and ultimately achievement of global grade-level standards.

Figure 2. Theory of Change

This theory is testable through three broad research questions:

1. **Inputs:** Did EdoBEST provide resources, build capacity, and institute quality assurance processes that enable change?
2. **Drivers**: Did EdoBEST inputs increase the amount of time for pupils to learn, create a more positive and safe learning environment, and raise instructional quality?
3. **Outcomes:** Did EdoBEST inputs and improvement across drivers translate into learning gains in English literacy and maths?

We measure inputs, drivers, and outcomes to test our theory of change and to track progress. This approach is important due to the very short length of the EdoBEST pilot. Typically, studies are powered to detect a minimum 0.2 standard deviation difference for an intervention lasting at least one year; the EdoBEST pilot lasted three months. We knew, given such a brief period, it may be difficult to detect learning impacts; however, we thought it valuable to assess whether the program was implemented with fidelity and, should that be the case, whether the strength of the implementation created the conditions that Bridge believes will lead to better pupil learning. Additionally, by testing English literacy and math skills now, we establish a baseline to benchmark future learning gains. Detailed explanation of our methodology can be found in *Appendix A*.

---

9 SUBEB selected the final list of schools to participate in the EdoBEST pilot. These schools generally had a sufficient number of classrooms and some desks in decent physical condition, offered Primary 1-4 classes, and had an adequate number of teachers. The selection process also ensured that each ward had at least one school represented in the pilot. In cases where no schools or multiple schools within a ward met the eligibility criteria, a ranked list of schools was established based on school population (serve the most children), facilities quality (most quickly begin piloting without construction interruptions), and number of teachers (provide as many teachers as possible with training and support).

10 Quality assurance includes monitoring and feedback processes, which generate data that enable continuous improvements in other inputs, such as instructional materials.

II. Data and Descriptive Statistics

The study includes 60 Edo primary schools (30 EdoBEST "treatment" and 30 status quo "control" schools).[11] We first sampled 30 EdoBEST schools with characteristics most representative of the 1,111 Edo State schools. Then, we applied an algorithm for each EdoBEST school to determine which non-participating ("status quo") school it most closely resembled. The algorithm worked well: Our data revealed no initial differences in observable characteristics between the 30 EdoBEST and 30 status quo schools.[12]

The assessment team conducted unannounced visits to the schools between July 16th to 30th, 2018. These visits included an interview with the headmaster, school observations, interviews with Primary 3 (P3) and Primary 4 (P4) teachers, classroom observations, and pupil assessments in maths and literacy. If a school had multiple classrooms in the same grade, we randomly selected one classroom to participate in the study.[13] *Figure 3* shows the final sample.

Figure 3. Sample Size by School Type[14]

Sample	EdoBEST	Status Quo
Schools Visited	30	30
P3 Classrooms Observed	30	30
P4 Classrooms Observed	29	30
P3 Teachers Surveyed	30	28
P4 Teachers Surveyed	26	28
P3 Pupils Assessed	883	784
P4 Pupils Assessed	836	735

Note: EdoBEST "schools" are those where at least one teacher (in addition to the head teacher) has participated in the EdoBEST program; however, the EdoBEST pilot covered only grades P1-P3. As such, P4 classrooms in EdoBEST schools did not formally undergo the EdoBEST program, though spillovers likely occurred.

A. Establishing a Fair Counterfactual: Using Non-Participating Primary 4 Classrooms to Understand Pre-Existing Differences Between EdoBEST and Status Quo Schools

To measure the efficacy of the EdoBEST program, we use a difference-in-difference method. We assessed inputs, drivers, and learning outcomes of P3 pupils across the sample of 30 EdoBEST and 30 status quo schools[15]. We also assessed P4 pupils in these same schools, using the same measurements. Since none of the P4 pupils have received the EdoBEST treatment, we expect their results to be similar across school types; should their results differ, we can use them to help us estimate any

pre-existing differences between pupils in the EdoBEST schools and the status quo schools. Note that for conciseness, we will refer to P4 in EdoBEST participating schools as "EdoBEST classrooms" or "EdoBEST teachers", but it is important to remember that they did not directly receive the EdoBEST intervention.

As expected, given EdoBEST was not rolled out to P4 classrooms, there were no differences in learning outcomes between P4 pupils in EdoBEST participating schools vs. status quo schools at the end of the year. There were also no differences in pupil characteristics.

Figure 4. Primary 4 Assessment Scores

	Maths			English Literacy		
	EdoBEST	Status Quo	Difference	EdoBEST	Status Quo	Difference
Pre-Primary 3 Level Questions (Revision)						
Mean	66.7%	68.1%	-1.4%	61.9%	62.5%	-0.6%
Median	76.9%	76.9%	0.0%	60.9%	60.9%	0.0%
Standard Deviation	(0.25)	(0.25)		(0.20)	(0.18)	
Primary 3 Level Questions (At Level)						
Mean	29.7%	31.1%	-1.4%	43.8%	43.2%	0.6%
Median	30.4%	30.4%	0.0%	42.9%	35.7%	7.2%
Standard Deviation	(0.13)	(0.13)		(0.25)	(0.23)	
Total						
Mean	43.1%	44.4%	-1.3%	56.6%	56.7%	-0.1%
Median	44.4%	47.2%	-2.8%	55.6%	55.6%	0.0%
Standard Deviation	(0.15)	(0.15)		(0.21)	(0.19)	

11 This powers the study to identify differences at roughly ⅜ standard deviations.

12 We used data from the Edo State School Census conducted in late 2017. We focused on variables that may affect pupil outcomes. Two broad types of variables were selected - those that showed the existing level of resources available at the school (teachers, facilities, etc.) and those that help capture the environment in which the schools operate (urban vs. rural). See Figure 30 in the Appendix A1 for the balance table.

13 All pupils within the participating classroom were administered the assessments.

14 Classroom observations were not conducted for one P4 teacher at one EdoBEST school. In this classroom, the P3 and P4 pupils were combined in one classroom and were taught by a teacher trained by the EdoBEST program. Teachers absent from the school on the day of the visit were not surveyed, resulting in a discrepancy in the number of teachers surveyed.

15 Detailed information on our measurement instruments can be found in *Appendix A2. Measurements.*

Figure 5. P4 Pupil Characteristics at Observed EdoBEST vs. Status Quo Classrooms

	EdoBEST	Status Quo	Difference
% Girls	50.4%	49.8%	0.6%
Median Age	11	11	0
% Overage	51.4%	59.2%	-7.8%

Note: Metrics are averages across observed Primary 4 classrooms. Pupils are classified as overage if they are 12 years old or over.

We also examined teacher-to-pupil and classroom-to-pupil ratios, inputs and drivers of learning, to see if EdoBEST participating schools were already better resourced or had existing schooling conditions more conducive to learning. The average EdoBEST P4 classroom had both more pupils and teachers (1.5 vs. 1.2 on average), resulting in a similar pupil-teacher ratio.

Figure 6. Average Number of Pupils and Teachers in Observed Primary 4 Classroom

	EdoBEST	Status Quo	Difference
Pupils Enrolled	**40.5**	**31.8**	**8.7**
Girls Enrolled	19.8	15.3	4.5
Boys Enrolled	20.7	16.5	4.2
Gender Parity Index	95.7%	92.7%	2.9%
Teachers Staffed	**1.5**	**1.2**	**0.3**
Pupil-Teacher Ratio	28.8	28.6	0.2

Note: Metrics are averages across observed Primary 4 classrooms.

Learning supports were roughly equivalent between EdoBEST and status quo P4 classrooms. These P4 classrooms had similar resources. The P4 teachers also received similar support within their schools, though more P4 teachers in EdoBEST schools attended training: on average, 1 in 5 teachers in EdoBEST schools attended a training in the past year, whereas 1 in 10 teachers in status quo schools did so.

Figure 7. Inputs to Support Learning in Primary 4 Classrooms

	EdoBEST	Status Quo	Difference
Teacher Support			
% Attended Training in Current School Year	15.4%	3.6%	11.8%
Average # of Trainings Attended	0.2	0.1	0.1
% Received Feedback from Outside Visitor in Past Month	53.8%	42.9%	10.9%
% Received Feedback from Head Teacher in Past Week	53.8%	50.0%	3.8%
% Using Electronic Device to Facilitate Lesson	0.0%	0.0%	0.0%
Learning Materials			
% of Classrooms with Textbooks Visible	58.6%	56.7%	1.9%
Average # of Books Accessible to Pupils	3.2	2.1	1.1

Note: Metrics are averages across observed Primary 4 classrooms. Electronic device usage and presence of textbooks was observed by assessors; all other metrics are self-reported by teachers to assessors.

Figure 8. Drivers of Learning in Primary 4 Classrooms

	EdoBEST	Status Quo	Difference
Time to Learn			
Observed Length of School Day (Hours)	6.5	5.7	0.8
% of Teachers Present in Morning	89.7%	93.3%	-3.6%
% of Teachers Present at 1:45 pm	86.7%	76.7%	10.0%
% of Teachers Present at 2:30 pm	53.3%	10.0%	43.3%
% of Pupils Present	75.9%	74.2%	1.7%
% of Classrooms with Either No Teacher or Pupils Present	10.3%	6.7%	3.6%
% of Classroom Observations Where …			
Pupils Were Focused on the Lesson for No or Few Minutes	18.4%	21.1%	-2.7%
Pupils Were Focused on the Lesson for Some Minutes	14.9%	8.9%	6.0%
Pupils Were Focused on the Lesson for Many or All Minutes	66.7%	70.0%	-3.3%
% of Classroom Observations Where …			
Many or All Minutes were Lost to Pupil Misbehavior	9.1%	6.1%	3.0%
Some Minutes were Lost to Pupil Misbehavior	3.9%	2.4%	1.5%
No or Few Minutes were Lost to Pupil Misbehavior	87.0%	91.5%	-4.5%
Classroom Environment			
% of Teachers Who Scolded, Belittled, or Used Corporal Punishment on Pupils	26.9%	42.9%	-16.0%
% of Teachers Who Scolded or Belittled	26.9%	42.9%	-16.0%
% of Teachers who Used Corporal Punishment	3.8%	21.4%	-17.6%
% of Teachers Who Praised or Positively Narrated Behaviour of Pupils	46.2%	46.4%	-0.2%
% of Teachers Who Praised Pupils	46.2%	46.4%	-0.2%
% of Teachers Who Praised at Least 1-2 Pupils	42.3%	46.4%	-4.1%
% of Teachers Who Praised at Least 3 Pupils	23.1%	25.0%	-1.9%
% of Teachers Who Praised the Whole Class	15.4%	7.1%	8.3%
% of Teachers Who Positively Narrated the Behavior of Pupils	3.8%	0.0%	3.8%
% of Teachers Who Narrated the Positive of Least 1-2 Pupils	3.8%	0.0%	3.8%
% of Teachers Who Narrated the Positive of Least 3 Pupils	0.0%	0.0%	0.0%
% of Teachers Who Narrated the Positive of the Whole Class	0.0%	0.0%	0.0%
Instructional Quality			
% of Classroom Observations Where …			
None or Few Pupils Followed Directions	3.9%	10.1%	-6.2%
Some Pupils Followed Directions	24.7%	6.3%	18.4%
Most or All Pupils Followed Directions	71.4%	83.5%	-12.1%
% of Classroom Observations with Independent Practice	35.1%	26.8%	8.3%
During Independent Practice, % of Classroom Observations Where …			
None or Few Pupils Working Hard	7.4%	13.6%	-6.2%
Some Pupils Working Hard	18.5%	0.0%	18.5%
Most or All Pupils Working Hard	74.1%	86.4%	-12.3%
Pupils Given Feedback			
% of Pupils Called On During Whole Class Time	16.3%	18.7%	-2.4%
% of Pupils Who Received Concrete Feedback in Front of Whole Class	10.0%	10.9%	-0.9%
% of Pupils Who Have Their Independent Work Checked	8.4%	13.6%	-5.2%

Note: Observations were conducted in three 10-minute segments per classroom. Each segment is treated as individual observation in these summary statistics.

When we look at teacher behaviors within their classrooms, two differences stand out that may point to the existence of positive spillovers. First, Primary 4 teachers at EdoBEST schools were five times as likely to still be at school at 2:30 pm, the official end of the school day under EdoBEST. Second, P4 EdoBEST teachers were less likely to have scolded, belittled, or used corporal punishment on pupils during the observed class (27% for EdoBEST versus 43% for status quo).

While the pattern of pupil behavior across these classrooms differed somewhat, overall pupilbehavior appears similar. EdoBEST had more classrooms where some pupils were following directions and were working hard some of the time, whereas status quo schools had more classrooms where more pupils followed directions all of the time.

B. Descriptive Differences: Comparing Primary 3 Classrooms in EdoBEST and Status Quo Classrooms

For P3, we see almost 7 more pupils in EdoBEST classrooms, but that appears to be offset by slightly more teachers as well, resulting in no differences in the pupil-teacher ratio. Pupils also look similar across the two school types.

Figure 9. Average Number of Pupils and Teachers in Observed Primary 3 Classroom

	EdoBEST	Status Quo	Difference
Pupils Enrolled	**34.7**	**31.8**	**2.9**
Girls Enrolled	16.8	14.2	2.6
Boys Enrolled	17.9	17.6	0.3
Gender Parity Index	93.9%	80.7%	13.2%
Teachers Staffed	**1.5**	**1.1**	**0.4**
Pupil-Teacher Ratio	25.3	28.2	-2.9

Note: Metrics are averages across observed Primary 3 classrooms.

Figure 10. P3 Pupil Characteristics at Observed EdoBEST vs Status Quo Classrooms

	EdoBEST	Status Quo	Difference
% Girls	47.7%	44.4%	3.3%
Median Age	10	10	0
% Overage	62.4%	58.9%	3.5%

Note: Metrics are averages across observed Primary 3 classrooms. Pupils are classified as overage if they are 11 years old or over.

1. Inputs

EdoBEST is designed to provide more teacher support. In practice, P3 EdoBEST classrooms are better supported: 97% of EdoBEST teachers attended training in the current school year, versus only 11% at status quo schools. They were also nearly twice as likely to receive feedback from an outside visitor.

The EdoBEST intervention was also designed to increase familiarity with information and communication technology (ICT) and leverage it to provide learning supports (particularly lesson guides for every class period), practices beyond the experiences of most Edo teachers. As expected, all EdoBEST teachers used electronic devices to facilitate lessons.

Figure 11. Inputs to Support Learning at P3 EdoBEST vs. Status Quo Classrooms

	EdoBEST	Status Quo	Difference
Teacher Support			
% Attended Training in Current School Year	**96.7%**	**11.1%**	**85.6%**
Average # of Trainings Attended	1.1	0.2	0.9
% Received Feedback from Outside Visitor in Past Month	80.0%	44.4%	35.6%
% Received Feedback from Head Teacher in Past Week	70.0%	66.7%	3.3%
% Using Electronic Device to Facilitate Lesson	**100.0%**	**0.0%**	**100.0%**
Learning Materials			
% of Classrooms with Textbooks Visible	**86.7%**	**53.3%**	**33.4%**
Average # of Books Accessible to Pupils	13.8	3.9	9.9

Note: Metrics are averages across observed Primary 3 classrooms. Electronic device usage and presence of textbooks was observed by assessors; all other metrics are self-reported by teachers to assessors.

Resources for pupils are also more evident in EdoBEST P3 classrooms. While textbooks were seen in just over half of status quo P3 classrooms, 87% of EdoBEST P3 classrooms had visible textbooks.

2. Drivers

We also observed differences among drivers of learning. EdoBEST pupils had more time to learn, experienced a better classroom environment, and received higher quality instruction.

Pupils appear to benefit from more instructional time, both through a longer school day (by almost one hour) and more time focused on lessons. Nearly all (94%) observations of EdoBEST classrooms saw most pupils focused on the lesson for the majority of the observed time, compared to 73% for the status quo P3 classrooms.

EdoBEST P3 classrooms also appear to have more positive learning environments. Praise is much more prevalent. While just under half of teachers in status quo classrooms praised pupils at least once during observations, nearly all of EdoBEST P3 teachers praised pupils. Furthermore, teachers were six times more likely to reinforce positive behavior by utilizing the "narrating the positive" technique, which points out a specific pupil action to encourage continuance and mirroring by peers. There may also be less negativity and intimidation in EdoBEST classrooms. Scolding, belittling, and/or corporal punishment occurred in 44% of status quo classrooms, but only in 23% of EdoBEST classrooms.

More quality instruction appears to be occurring as well. EdoBEST classrooms were more often characterized by pupils following directions, pupils working hard, and teachers providing feedback to pupils during independent practice time.

A well-run classroom will have all pupils following the teacher's directions. In EdoBEST classrooms, most pupils followed directions in 93% of observations, compared to 82% in status quo classrooms. EdoBEST pupils exhibited greater focus, too, with most pupils focused on the entire lesson and few minutes where pupils were unfocused. Nearly 18% of observations for status quo classrooms showed that pupils were unfocused throughout the lesson.

Pupils also need the opportunity to engage in a productive struggle, applying what they just learned. This requires independent practice time. 61% of EdoBEST classroom observations contained independent practice time, compared to 43% in status quo schools. Pupils in EdoBEST classrooms were also more likely to be working hard during this time; during nearly all independent practice time, EdoBEST pupils were concentrating on understanding text or solving problems.

Finally, receiving clear feedback is critical to pupils' growth. A healthy learning process, Bridge posits, ensures pupils know whether their answer is correct, and if not, why and how to correct it. As such, a successful independent practice will not just have pupils working hard, but will also have a teacher circulating and giving one-on-one feedback to each pupil. EdoBEST pupils were more than three times more likely to have their work checked by a teacher during independent practice.

Figure 12. Drivers of Learning in P3 EdoBEST vs. Status Quo Classrooms

	EdoBEST	Status Quo	Difference
Time to Learn			
Observed Length of School Day (Hours)	6.5	5.7	0.8
% of Teachers Present in Morning	100.0%	93.3%	6.7%
% of Teachers Present at 1:45 pm	100.0%	73.3%	26.7%
% of Teachers Present at 2:30 pm	90.0%	13.3%	76.7%
% of Pupils Present	81.9%	71.8%	10.1%
% of Classrooms with Either No Teacher or Pupils Present	0.0%	10.0%	-10.0%
% of Classroom Observations Where …			
Pupils Were Focused on the Lesson for No or Few Minutes	2.2%	17.8%	-15.6%
Pupils Were Focused on the Lesson for Some Minutes	3.3%	8.9%	-5.6%
Pupils Were Focused on the Lesson for Many or All Minutes	94.4%	73.3%	21.1%
% of Classroom Observations Where …			
Many or All Minutes were Lost to Pupil Misbehavior	2.2%	1.2%	1.0%
Some Minutes were Lost to Pupil Misbehavior	3.3%	2.5%	0.8%
No or Few Minutes were Lost to Pupil Misbehavior	94.4%	96.3%	-1.9%
Classroom Environment			
% of Teachers Who Scolded, Belittled, or Used Corporal Punishment on Pupils	23.3%	44.4%	-21.1%
% of Teachers Who Scolded or Belittled	23.3%	37.0%	-13.7%
% of Teachers who Used Corporal Punishment	6.7%	14.8%	-8.1%
% of Teachers Who Praised or Positively Narrated Behaviour of Pupils	96.7%	48.1%	48.6%
% of Teachers Who Praised Pupils	96.7%	44.4%	52.3%
% of Teachers Who Praised at Least 1-2 Pupils	73.3%	44.4%	28.9%
% of Teachers Who Praised at Least 3 Pupils	53.3%	22.2%	31.1%
% of Teachers Who Praised the Whole Class	83.3%	11.1%	72.2%
% of Teachers Who Positively Narrated the Behavior of Pupils	46.7%	7.4%	39.3%
% of Teachers Who Narrated the Positive of Least 1-2 Pupils	36.7%	7.4%	29.3%
% of Teachers Who Narrated the Positive of Least 3 Pupils	16.7%	0.0%	16.7%
% of Teachers Who Narrated the Positive of the Whole Class	26.7%	0.0%	26.7%
Instructional Quality			
% of Classroom Observations Where …			
None or Few Pupils Followed Directions	0.0%	7.4%	-7.4%
Some Pupils Followed Directions	6.7%	11.1%	-4.4%
Most or All Pupils Followed Directions	93.3%	81.5%	11.8%
% of Classroom Observations with Independent Practice	61.1%	43.2%	17.9%
During Independent Practice, % of Classroom Observations Where …			
None or Few Pupils Working Hard	0.0%	11.4%	-11.4%
Some Pupils Working Hard	1.8%	25.7%	-23.9%
Most or All Pupils Working Hard	98.2%	62.9%	35.3%
Pupils Given Feedback			
% of Pupils Called On During Whole Class Time	21.6%	17.0%	4.6%
% of Pupils Who Received Concrete Feedback in Front of Whole Class	15.3%	10.5%	4.8%
% of Pupils Who Have Their Independent Work Checked	32.0%	9.8%	22.2%

Note: Observations were conducted in three 10-minute segments per classroom. Each segment is treated as individual observation in these summary statistics.

3. Outcomes

EdoBEST pupils scored higher on every exam. The questions for each subject exam can be broken down into "revision", which are questions based on topics that should have been covered prior to P3, and "at-level", which are questions based on topics that should be covered during P3.

**Figure 13. P3 Assessment Scores at EdoBEST vs. Status Quo Classrooms**

	Maths			English Literacy		
	EdoBEST	Status Quo	Difference	EdoBEST	Status Quo	Difference
Pre-Primary 3 Level Questions (Revision)						
Mean	63.6%	57.7%	5.9%	59.0%	52.8%	6.2%
Median	69.2%	61.5%	7.7%	56.5%	52.2%	4.3%
Standard Deviation	(0.29)	(0.28)		(0.17)	(0.18)	
Primary 3 Level Questions (At-Level)						
Mean	26.7%	26.3%	0.4%	37.3%	34.4%	2.9%
Median	26.1%	26.1%	0.0%	35.7%	28.6%	7.1%
Standard Deviation	(0.13)	(0.12)		(0.21)	(0.20)	
Total						
Mean	40.0%	37.7%	2.3%	52.2%	47.1%	5.1%
Median	41.7%	38.9%	2.8%	50.0%	44.4%	5.6%
Standard Deviation	(0.16)	(0.15)		(0.17)	(0.17)	

On the maths exam, EdoBEST pupils scored 2 percentage points higher than status quo pupils. When we break the test down by difficulty, EdoBEST pupils scored about 6 percentage points higher on the revision section, which is about one extra question on this 13-question portion of the test. For at-level content, there is essentially no difference. Since the content portion of the exam consists solely of five-choice multiple choice questions, a 26% accuracy rate is only slightly higher than the score produced by selecting random responses, suggesting that pupils across both school types were likely guessing for many questions. This indicates that the P3 Edo curriculum, without more time for remedial education, is beyond what prior schooling has prepared the average Edo P3 pupil to capably grasp.

On English literacy, EdoBEST pupils scored 5 percentage points higher than status quo pupils. This difference was driven by their performance on revision items. On grade level content, pupils across both school types were able to answer approximately a third of the questions correctly, which is better than guessing given the four-choice multiple choice question exam. However, since the test is focused on literacy itself rather than on at-level grammar or vocabulary, the overall low performance suggests pupils are struggling here as well.

Finally, the EdoBEST program appears to impact pupils at all levels of academic performance. While performance on at-level content questions does not differ between the two groups, EdoBEST pupils outperform status quo pupils at every point in the score distribution. This means star, average, and struggling pupils alike all benefit from EdoBEST.

Figure 14. Distributions of P3 Assessment Scores at EdoBEST vs. Status Quo Classrooms
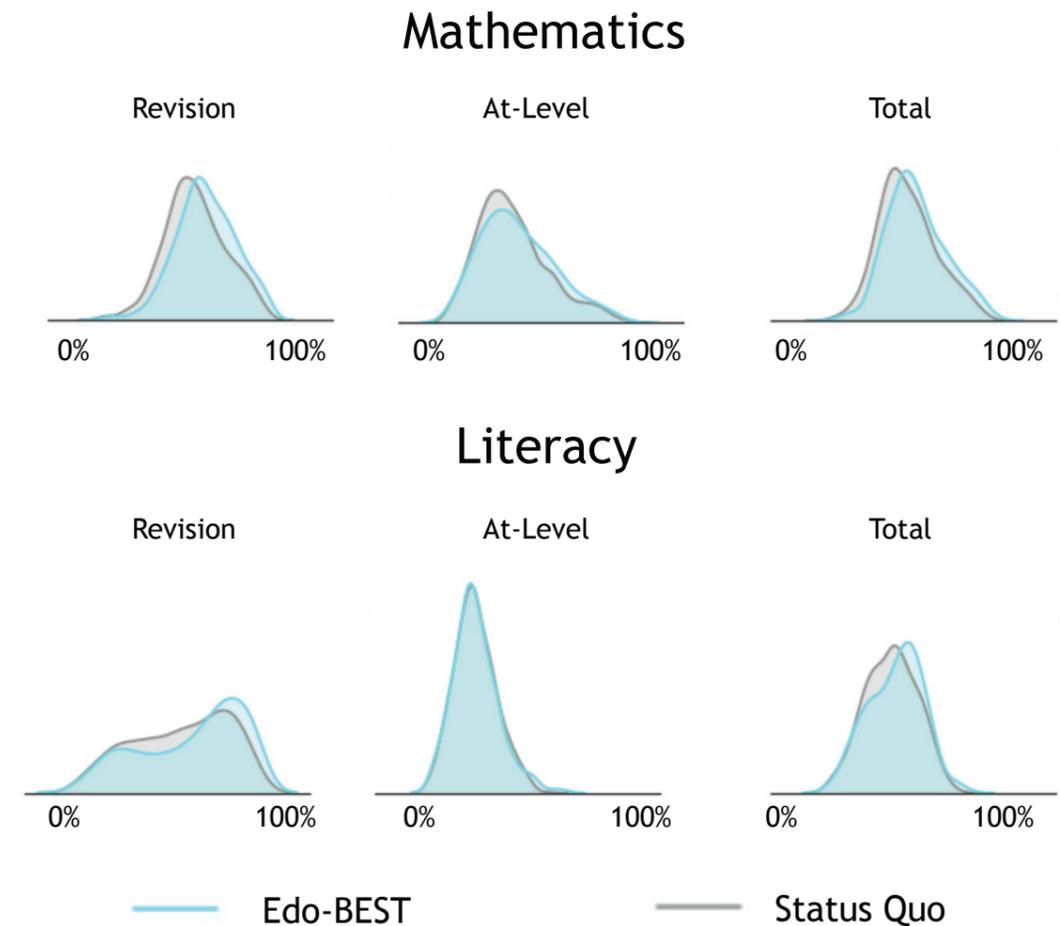
III. Results

A. A Controlled Study: The Counterfactual and Differences in Other Non-Intervention Characteristics

While the above summary statistics provide good indications of the differences between EdoBEST and status quo inputs, drivers of learning, and pupil learning achievement, we can more rigorously isolate the impact of EdoBEST by using a regression framework to take into account non-intervention variables that might affect outcomes. We account for geographic location, class size, and the number of additional teachers in the classroom. For learning outcomes, we also include pupil age and gender.

We also account for differences across schools that are not the direct result of EdoBEST. A school may be the recipient of a pre-existing benefit that positively affects teaching or pupil performance. By including results for P4 classrooms, none of which received the EdoBEST intervention, we separate out the effects of any pre-existing programs and ensure they do not distort our measure of the EdoBEST impact.

For example, imagine an incredibly capable headmaster at an EdoBEST school. This headmasters' influence will likely help improve learning outcomes for both P3 and P4 pupils. If we only compare P3 outcomes between EdoBEST and status quo schools, our estimates of EdoBEST's impact may be erroneously large, as they will also contain the effect of this excellent headmaster. This headmaster effect, however, will apply both to P3 and P4 classrooms in the EdoBEST school. Taking the *difference* between the two grade levels therefore first eliminates the headmaster effect. Then, comparing this difference with that of status quo schools leaves us with the impact of EdoBEST. In short, the difference-in-differences method will make proper adjustments to allow for a fair comparison.[16]

This approach, however, also gives conservative estimates of the EdoBEST effect in cases where the P4 classroom has indirectly benefited from the EdoBEST intervention on other grades and the headmaster. For example, it is possible that P4 teachers at EdoBEST schools stay longer due to peer pressure or inspiration from their P3 colleagues following the EdoBEST schedule. Or, since the headmaster was also trained in EdoBEST, s/he may manage the entire school in the manner prescribed by the program. For example, we see fewer instances of corporal punishment in EdoBEST schools, likely attributable to EdoBEST training. In these cases, the positive, additional consequences of EdoBEST is not only uncaptured, it actually reduces the EdoBEST impact estimate by mistaking it for a pre-existing program. We refer to these additional benefits as spillovers.[17]

---

16 In technical terms, we apply this approach by analyzing both Primary 3 and Primary 4 results together, using interacted terms for both school type (with EdoBEST vs status quo) and grade. If the descriptive differences between EdoBEST and status quo P3 classrooms are driven by unobserved school level differences rather than by the actual EdoBEST intervention, the entire effect will be captured by the EdoBEST school dummy variable, while the EdoBEST-P3 interaction term will be insignificant. All regression models use these interacted terms.

17 Spillovers are likely to occur within EdoBEST participating schools. Channels include school leadership (headmasters of EdoBEST participating schools were also trained in EdoBEST); across teachers (from Primary 3, with an EdoBEST trained teacher, to Primary 4 classrooms, without an EdoBEST trained teacher); and/or across siblings (pupils in P4 may a sibling in P3).

B. Learning Resources Have Increased Substantially

The EdoBEST program has dramatically increased teacher support and pupil learning materials. EdoBEST teachers were 73% more likely to have attended teacher training in the current school year[18]. EdoBEST teachers were the only ones who used an electronic device to facilitate teaching, consistent with the lack of technology use overall in Edo schools. On average, there were 9 more textbooks visible and accessible to pupils in EdoBEST classrooms.

We are unable to detect an impact on the frequency of on-going coaching. Although EdoBEST classrooms were 36% more likely to receive feedback from an outside visitor, once we account for other factors, the adjusted difference (25%) is no longer statistically significant.

Figure 15. More Teacher Support, More Learning Materials

Learning Supports		EdoBEST Effect	
Teacher Support			
(1)	% of Teachers Attended Training	73%	**
(2)	# of Trainings Attended	0.76	**
(3)	% of Teachers Received Outsider Feedback	25%	
(4)	% of Teachers Received Headmaster Feedback	-7%	
(5)	% of Teachers Used Electronic Device	100%	**
Learning Materials			
(6)	% of Classrooms w/ Textbooks Visible	33%	+
(7)	# of Textbooks Accessible to Pupils	8.98	*

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (+) at the 90% confidence level.

Figure 16. Learning Resources in EdoBEST Classrooms vs. the Status Quo

Full regression results can be seen in Appendix B1. Inputs to Support Learning.

C. Drivers of Learning Have Improved

1. More Time for Learning

EdoBEST provided more time where teachers were present and pupils were focused on the lesson. Teachers at EdoBEST were more likely be at school in the afternoon. On average, they remained for 45 minutes longer at EdoBEST schools, and P3 EdoBEST teachers were 38% more likely to still be at school at 2:30 PM.[19]

Figure 17. More Time for Learning

Driver Metric		EdoBEST Effect	
Teachers & Pupils Present			
(1)	Hours in the School Day*	0.75	**
(2)	% of Teachers Present in Morning	9%	
(3)	% of Teachers Present at 1:45 PM	19%	
(4)	% of Teachers Present at 2:30 PM	38%	**
(5)	% of Pupils in Attendance	8%	
(6)	% of Classrooms with Either No Teacher or Pupils Present	-13%	
More Instructional Time			
(7)	Time Pupils are Focused on Lesson	0.33	+
(8)	Time Lost to Pupil Misbehavior	-0.05	

Note: (7) and (8) are best interpreted as changes of the phenomenon, where the following rating scale is used: -1 for no/few minutes, 0 for some minutes, and 1 for many/all minutes. Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (+) at the 90% confidence level.

We also examine the degree to which pupils had focused time in the classroom. Pupils were highly focused in nearly all (94%) EdoBEST classrooms.

Figure 18. Minutes Where Most Pupils Are Focused on Lesson
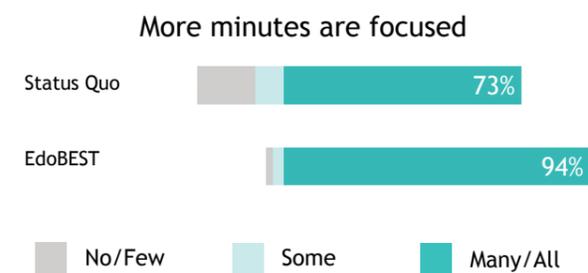
2. A Better Learning Environment

EdoBEST created a more positive learning environment. P3 EdoBEST teachers were more likely to use praise by 54%. This is not just limited to generic praise ("Good job!", "Well done!"); P3 EdoBEST teachers were 35% more likely to identify and narrate specific, positive behaviors of pupils in the classroom ("Eno is sitting tall!"). Furthermore, teachers were 18% more likely to distribute praise to multiple pupils within the classroom, rather than isolating praise on one or two "star" pupils.

Figure 19. A Better Learning Environment

Driver Metric		EdoBEST Effect	
Punishment & Negative Culture			
(1)	% of Teachers Who Scolded or Belittled	6%	
(2)	% of Teachers who Used Corporal Punishment	7%	
Praise & Positive Discipline			
(3)	% of Teachers Who Praised Pupils	54%	**
(4)	% of Teachers Who Praised at Least 1-2 Pupils	33%	*
(5)	% of Teachers Who Praised at Least 3 Pupils	34%	*
(6)	% of Teachers Who Praised the Whole Class Together	65%	**
(7)	% of Teachers Who Positively Narrated the Behavior of Pupils	35%	**
(8)	% of Teachers Who Narrated the Positive for 1-2 Pupils	26%	*
(9)	% of Teachers Who Narrated the Positive for 3 or More Pupils	18%	*
(10)	% of Teachers Who Narrated the Positive to the Whole Class Together	28%	**

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (+) at the 90% confidence level.

Surprisingly, the analysis shows no differences on punishment as a common practice once other factors are considered. We believe this is due to a methodological limitation rather than a reflection of on-the-ground truth, given the large differences in teacher behavior and our inability to capture spillovers. Three times as many teachers at status quo schools used corporal punishment compared to teachers at EdoBEST schools. This is consistent with the literature that demonstrates that a positive school climate and effective leadership are linked to positive educational outcomes.[20] EdoBEST headmasters, in the service of pupil welfare, were likely to apply these new expectations of restraint across the entire school, without regard to whether teachers participated in EdoBEST.

3. Improved Instructional Quality

EdoBEST pupils experienced better quality instruction in all three measured categories of instructional quality.

Figure 20. Pupils are More Likely to Follow Directions, Work Hard, and Get Feedback

Driver Metric		EdoBEST Effect	
Pupils Follow Directions			
(1)	Pupils Follow Directions	0.27	+
Pupils Work Hard Independently			
(2)	% Obs. Where Independent Practice Occurred	11%	
(3)	Pupils Work Hard During Independent Practice	0.56	*
Pupils Answer Questions & Get Feedback			
(4)	% Pupils Called On in Front of Whole Class	7%	
(5)	% of Pupils Who Received Concrete Feedback in Front of Whole Class	6%	
(6)	% Pupils Who Have Work Checked During Independent Practice	27%	**

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (+) at the 90% confidence level.

More pupils following the teacher's directions is an indication that the classroom is better-run. EdoBEST classrooms were more likely to have pupils follow directions: on a scale with -1 being no/few pupils are following directions and 1 being many/all pupils follow directions, EdoBEST classrooms scored 0.92 compared to 0.75 for status quo classrooms.[21]

Pupils in EdoBEST classrooms were also more likely to engage in the productive struggle to learn and master content. More pupils in EdoBEST classrooms worked hard during independent practice.[22] Again on a scale of -1 (no/few pupils worked hard) to 1 (many/all pupils worked hard), EdoBEST pupils scored 1.0 compared to 0.45 by their peers at status quo. Most notably, pupils were 27% more likely to receive clear feedback from teachers during this practice time in EdoBEST classrooms than in status quo P3 classrooms.

Figure 21 below illustrates the differences between EdoBEST and status quo P3 classrooms on the two instructional quality metrics as percent breakdowns by category. On both measures, a significantly larger portion of EdoBEST observations are reported in the "many/all" category.

---

20 "Relationships Between Measures of Leadership and School Climate," Education.

21 Values were assigned based on assessor observations as follows: no/few pupils followed directions (-1), some pupils followed directions (0), or many/all pupils followed directions (1).

22 Values were assigned based on assessor observations as follows: no/few pupils work hard (-1), some pupils work hard (0), or many/all pupils work hard (1).

Figure 21. EdoBEST Classrooms are More Productive

More pupils follow directions

Status Quo	81%
EdoBEST	93%

More pupils work hard

Status Quo	63%
EdoBEST	98%

■ No/Few ■ Some ■ Many/All
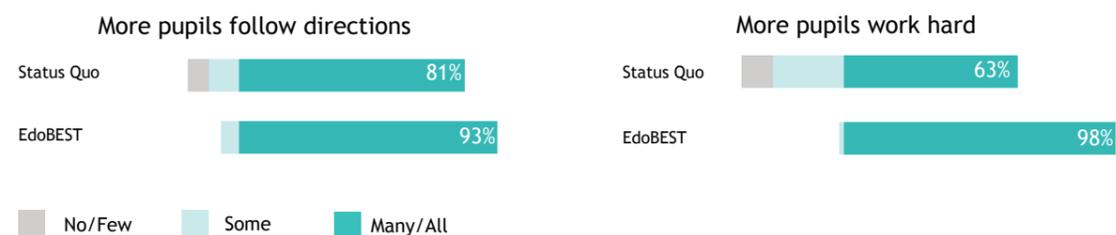
Full regression results for the EdoBEST's impact on teacher and pupil behavior can be seen in *Appendix B2. Drivers of Learning.*

D. Learning is Happening

Despite the short length of the pilot, we find that EdoBEST has already had a significant impact on maths and English literacy revision performance. P3 EdoBEST pupils scored 7 and 6 percentage points higher on maths revision and literacy revision respectively, which translates into an effect size of 0.27 and .33 standard deviations (see *Figure 26)*. While these results are promising, pupils are still struggling to master at-grade level content. Further remedial work is likely necessary to establish a stronger literacy and numeracy foundation before pupils can succeed with the national curriculum.[23]

_____

23 For the full assessment regression results, with both raw (% questions correct) and standardized scores, see Appendix *B3. Learning Achievement.*

On average, the EdoBEST program resulted in a 7% increase on the 13-item revision portion of the mathematics test, which contains content that pupils must master before moving on to at-level and more advanced concepts. Relative to the performance of P4 pupils, EdoBEST pupils have almost closed the gap *between* grades in less than three months. The average difference in one grade level, between P3 pupils and P4 pupils, was 10%.[24] In improving by 7%, P3 EdoBEST pupils attained revision scores closer to those of P4 pupils than those of their P3 peers in status quo classrooms, as can be seen in *Figure 22* below.

Figure 22. Maths Revision and At-Level Performance

Maths: Revision

Status Quo	Edo-BEST	Status Quo	Edo-BEST
58%	64%	68%	67%
Primary 3		Primary 4	

Maths: At Level

Status Quo	Edo-BEST	Status Quo	Edo-BEST
27%	26%	32%	29%
Primary 3		Primary 4	
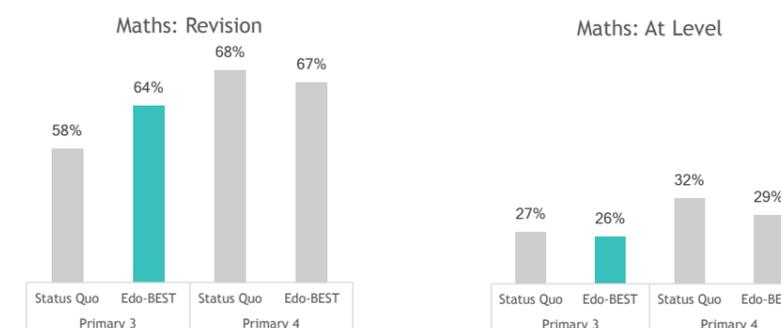
We turn to individual question-level analysis to examine how EdoBEST P3 pupils are closing this gap. For example, on question 29 (subtracting a one-digit number from a two-digit number without carrying), we find that 61% of P3 pupils in EdoBEST schools answered the question correctly compared to 55% of their status quo P3 peers. This increase of 6% is roughly halfway to the P4 status quo results. This pattern holds true across maths revision. Unfortunately, even as EdoBEST pupils outperform the status quo, these results illustrate that roughly 1/3 of these pupils still struggle with basic addition and subtraction. The results for the revision addition and subtraction questions can be seen in *Figure 23* below.

_____

24 This is after controlling for confounding factors.

Figure 23. Item-Level Performance on Maths Revision

19. How many apples?

75% 75% 85% 82%
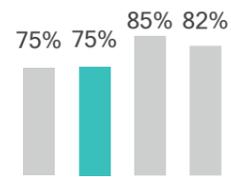
20. Which is more, 3 or 7?

70% 76% 77% 78%
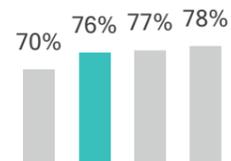
21. Compare the numbers: 6 _____ 5

41% 60% 44% 50%
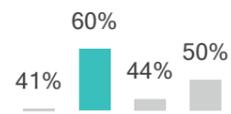
22. Compare the numbers: 17 _____ 23

41% 61% 46% 44%

SQ EB SQ EB
Primary 3 Primary 4

23. 4 + 5 = _____

69% 70% 76% 77%

24. 10 − 7 = _____

49% 62% 60% 55%

25. 8 + 8 = _____

63% 66% 77% 74%

26. 15 − 7 = _____

63% 67% 75% 75%

27. 11 + 8 = _____

64% 69% 77% 75%

SQ EB SQ EB
Primary 3 Primary 4

28. 22 + 6 = _____

67% 71% 79% 79%

29. 36 − 4 = _____

55% 61% 66% 63%

30. 15 + 9 = _____

72% 75% 82% 80%

31. 43 − 7 = _____

53% 56% 65% 60%

SQ EB SQ EB
Primary 3 Primary 4
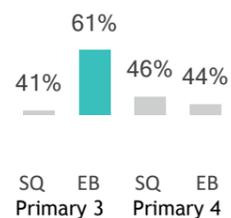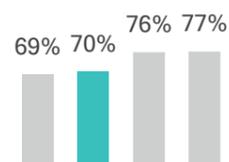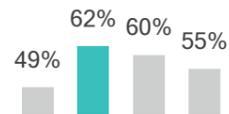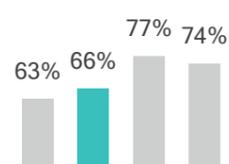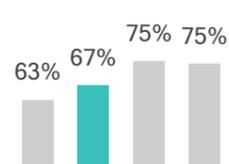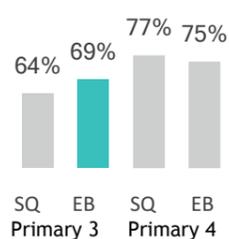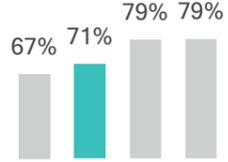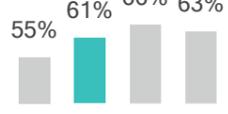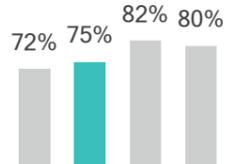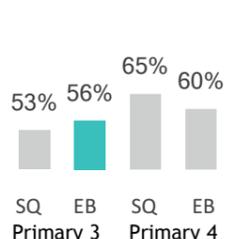
2. Literacy

On average, the EdoBEST program resulted in a 6% increase on the 22-item literacy revision test. Relative to the performance of P4 pupils on the same exam, EdoBEST pupils have started to close the gap *between* grades in less than three months. The average difference between P3 and P4 pupils, after controlling for confounding factors, was 11%.

Figure 24. Literacy Revision and At-Level Performance

Literacy: Revision

53% 57% 64% 62%

Status Quo | EdoBEST | Status Quo | EdoBEST
Primary 3 | Primary 4

Literacy: At Level

35% 36% 44% 44%

Status Quo | EdoBEST | Status Quo | EdoBEST
Primary 3 | Primary 4
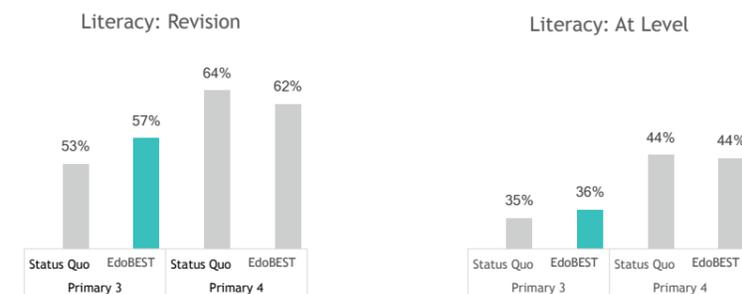
The literacy revision exam covered basic reading comprehension skills that P3 and P4 pupils should have already mastered. Looking at questions on a simple reading passage, we find that EdoBEST pupils were more likely to correctly answer comprehension questions. Once again, we see that a large portion of EdoBEST pupils still struggled on the most basic questions, and the vast majority could not answer the slightly more difficult question.

Figure 25. Performance on Basic Reading Comprehension

I am Pam. I see my dad. My dad can cook. I can watch. We can eat!

20. Who is in the story?

58% 67% 70% 75%

P3 Status Quo | P3 EdoBEST | P4 Control | P4 EdoBEST
Primary 3 | Primary 4

21. What can dad do?

60% 61% 58% 60%

P3 Status Quo | P3 EdoBEST | P4 Control | P4 EdoBEST
Primary 3 | Primary 4

22. What can they both do?

17% 24% 29% 30%

P3 Status Quo | P3 EdoBEST | P4 Control | P4 EdoBEST
Primary 3 | Primary 4
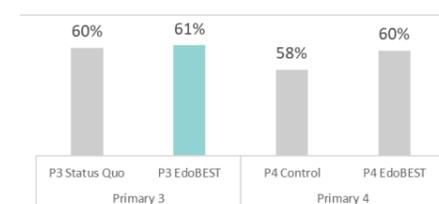
3. Translating Effect Sizes into % of Additional Schooling

Figure 26 below shows the impact of EdoBEST on each component of the assessment results, both in % questions correct and standard deviations (effect size). The gap between the status quo P3 and P4 average standardized score is also shown, providing an indication of how much status quo pupils grow during a typical school year. Consequently, we can convert the EdoBEST impact into a percent of a typical school year's learning, as can be seen in the final column of each subject.

Figure 26. Impact of EdoBEST on Assessment Scores

|	Math					English Literacy				
	EdoBEST Effect		P3 vs. P4 Gap		EdoBEST Effect as % of Gap	EdoBEST Effect		P3 vs. P4 Gap		EdoBEST Effect as % of Gap
	In % Correct	In SDs	In % Correct	In SDs		In % Correct	In SDs	In % Correct	In SDs	
Revision	7%*	0.27*	10%	0.38	71%	6%*	0.33*	11%	0.61	54%
At-Level	2%	0.12	5%	0.35	34%	2%	0.08	10%	0.45	18%
Total	4%+	0.24+	7%	0.43	56%	5%	0.25	11%	0.61	41%

Note: The P3 vs P4 gap is calculated as the mean difference in standardized scores between P3 and P4 pupils in status quo schools.

EdoBEST impacts on revision results translate into sizeable portions of typical annual learning. The typical gap on revision content between grades P3 and P4 is equivalent to about 0.38 standard deviations for math and 0.61 for English literacy. The EdoBEST program resulted in an additional 0.27 and 0.33 standard deviations of learning for math and literacy, respectively. By dividing the impacts by typical annual learning, we can interpret the EdoBEST impact from a single term as roughly 71% more annual schooling on math revision topics and 54% more on literacy revision topics.

Figure 27. Projections for Mastery of Revision Material

Projection for Maths Revision

Projection for Literacy Revision
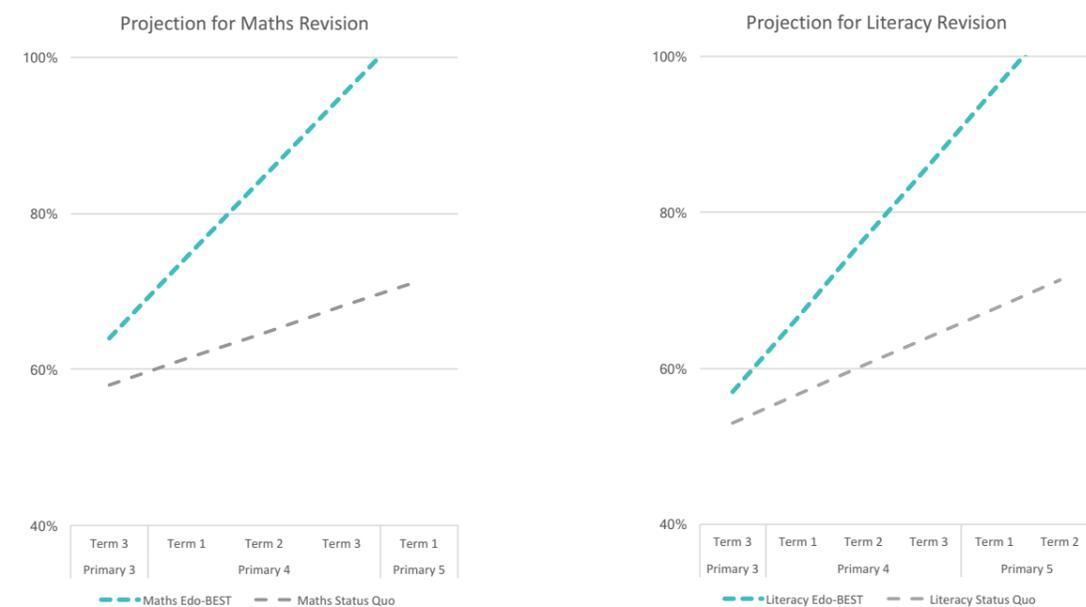
A 6-7% test score gain being equivalent to 54-71% more annual schooling underscores a fundamental problem and the vast challenge ahead: Edo state pupils are several grade levels behind on the national curriculum. Assuming EdoBEST can deliver learning gains at its current pace, the average P3 pupil will master *revision* content after four terms of instruction[25]. This means the EdoBEST P3 pupils will master P2-level or earlier content by the time they are in P5, rather than P7, putting them 3 rather than 5 years behind. If EdoBEST can adapt programming and double gains, mastery of foundational skills could happen in less than a year.

---

25 Assuming linear gains, EdoBEST P3 pupils will improve on revision level maths material by 3.33% a term (the average Edo gains) plus an additional 7% for a total of 10.33% per term. By the end of Primary 4, average performance would reach 95%. For literacy revision, EdoBEST P3 pupils will improve by 3.67% a term (the average Edo gains) plus an additional 6% for a total of 9.67% per term. By the end of Term 1 in Primary 5, average performance would reach 96%.

Girls outperformed boys across both subjects at all schools, but the gains for girls in EdoBEST classrooms built significantly on this. On the literacy test, the average P3 EdoBEST girl answered 6% more questions correct. On the math test, average total scores were higher by 4%. Additional gains come from improvements on both revision and at-level material.

While we are heartened by the performance of EdoBEST girls, particularly in the face of global inequities disadvantaging girls, we are also concerned that the benefits of EdoBEST have not fully reached boys. As EdoBEST is anticipated to reach more pupils in the coming years, it would be worthwhile to exploring and understanding why.

Figure 29. P3 Average Assessment Scores, Girls vs. Boys

IV. Girls are Driving Results in Learning Outcomes

We also want to know whether or not EdoBEST has made a difference in learning outcomes equally for girls and boys.[26] Remarkably, the impact of EdoBEST on learning appears to be driven entirely by girls. In fact, girls in EdoBEST classrooms are learning more than their status quo peers in both subjects, on both revision and at-level questions. Full regression results can be seen in *Appendix B4. Heterogeneous Impact.*

Figure 28. Impact of EdoBEST on Assessment Scores for Girls

	Math		English Literacy	
	% Correct	Effect Size	% Correct	Effect Size
Revision	7%**	0.27**	5%**	0.26**
At-Level	2%*	0.16*	7%**	0.32**
Total	4%**	0.26**	6%**	0.32**

---

26 We use regressions that isolate the differential maths and English literacy performance of girls in Primary 3 EdoBEST classrooms. These regressions contain the same variables and controls included in the overall learning gains regressions, with the addition of a gender binary term that interacts with classroom type. This interaction creates a specific variable for the EdoBEST-P3-girl combination.

V. Discussion

A. Study Limitations

We face four broad limitations in our study: power, selection bias, generalizability, and spillover effects. First, due to sample constraints, we are only able to detect large differences between the EdoBEST program and status quo schools. We may have missed less significant impacts on pupil performance. For our study, the effect has to be larger than ⅜ of a standard deviation for us to detect it; most research studies can detect impacts about half this size, anything greater than ▫ of a standard deviation.

Second, EdoBEST schools were chosen by SUBEB based on specific observable characteristics, and teacher participation was voluntary; if selected schools and participating teachers are systematically different, then EdoBEST schools may be on different learning trajectories and our estimates of the program could be biased. The fact that P4 classrooms look similar in EdoBEST and status quo schools (except for the incidence of corporal punishment, which may be the result of new school management practices by EdoBEST headmasters) suggests this selection bias is minimal.

Third, if selected schools within our sample are different than the average Edo state school our results would not be generalizable to a larger EdoBEST program roll-out. We improved generalizability through our sampling process, selecting the schools that were most representative of Edo state based on observable characteristics. Schools in our study look more similar to the average Edo state school than the average EdoBEST school as seen in *Appendix A1. Sample*.

Finally, the EdoBEST pilot was implemented in Primary 1 through Primary 3, but it is possible that there were positive spillover effects to Primary 4 pupils in EdoBEST schools. This would underestimate the EdoBEST impact, because the difference-in-differences method treats positive spillovers as pre-existing conditions. While it is impossible to completely control for these limitations, our study design should mitigate most of these effects. A more detailed explanation of these limitations can be found in *Appendix A4. Limitations Explained*.

B. Lessons Learned

EdoBEST is an ambitious program that aims to transform the education sector in Edo and provide a better-quality education in a state that has traditionally struggled to provide high quality education. We measured EdoBEST's achievements during its first term. Because of the short duration of the intervention, we focused our measurements on identifying how EdoBEST impacted inputs that support learning as well as the various drivers of learning (time, environment, instructional quality).

The results thus far indicate that EdoBEST has been successful in delivering inputs to support learning. EdoBEST teachers underwent professional training, used technology in the classroom, and had more textbooks accessible in the classroom. These inputs were utilized in a transformed classroom. Teachers behaved differently: they were at school longer, facilitated more productive independent practice,

checked pupils work more often, and used more praise in the classroom. EdoBEST pupils also behave differently: they work harder, are more focused, and are more likely to follow directions. A greater opportunity to learn has already yielded gains on pupil performance in maths and English literacy. In both subjects, the impact of EdoBEST on revision content in 3-months is equivalent to more than 60% of what is learned in a typical P4 school year.

Even as the EdoBEST program has improved learning outcomes and built a solid foundation on which to continue teacher and pupil growth, we recognize the need for more progress. Not all of the core EdoBEST inputs reached every classroom during the pilot. On teacher support, 10% of EdoBEST teachers were not visited by their headmaster in the past week. Similarly, 10% of teachers did not report an outside observer in the past month. On material support, 13% of EdoBEST classrooms had no visible textbooks.

We also see that EdoBEST's innovative teaching practices have not yet fully replaced existing practices. During 7% of our EdoBEST classroom observations, we witnessed corporal punishment, and minutes are still being lost to misbehavior. There is an opportunity to increase learning time with more productive discipline. There is also room for teachers to provide more feedback to more pupils. With only 15% of observed pupils speaking and hearing feedback in front of the class, and only 32% of pupils receiving feedback on independent work, improvements here have the potential to transform learning.

Finally, these results suggest the need to rethink curriculum in Edo State. While high expectations for all pupils are admirable and even necessary, the system has left the vast majority of pupils underprepared. Pupils are generally guessing on math and questions "at-level" with the state curriculum, which suggests they lack basic skills. Repair is needed. This may be the time to revisit the state curriculum and allow schools to meet pupils where they are at, then accelerate their learning to catch up to grade level expectations on the earliest reasonable timeline. Findings from this report can support teaching at the right level. While this remains speculative, it is possible that, given a full year of instruction, the P3 pupils in EdoBEST would have surpassed the performance of P4 pupils. This possibility means that Edo State may not need to compromise: Education policies can maintain high expectations, make allowances for pupils' current learning levels, and expect an accelerated student learning trajectory to bring more pupils to grade level faster.

Works Cited

Distribution of Population in National Assembly, Nigeria 2012. September 25, 2015. Distributed by Nigeria Data Portal. http://nigeria.opendataforafrica.org/apps/atlas/Edo/Total-Population.

Editors of Encyclopedia Britannica. "Edo." Britannica.
https://www.britannica.com/place/Edo-state-Nigeria.

"Edo Plans Tech-based Teaching, Learning in Public Schools - Obaseki." *Vanguard*, January 30, 2018.

"Edo Government Constitutes State Universal Basic Education Board (SUBEB)." Edo SUBEB (blog), August 20, 2018.

Majgaard, Kirsten and Alain Mingat. 2012. Education in Sub-Saharan Africa: A Comparative Analysis. Washington, D.C.: World Bank. DOI: 10.1596/978-0-8213-8889-1.

MDG Report 2015: Assessing Progress in Africa toward the Millennium Development Goal. Addis Ababa: ECA Printing and Publishing Unit, 2015.

National Population Commission (Nigeria) and RTI International, 2011. Nigeria Demographic and Health Survey (DHS) EdData Profile 1990, 2003, and 2008: Education Data for Decision-Making.

2011. Washington, DC, USA:National Population Commission and RTI International.

Onyekwelu, Stephen. "Nigerian Children to Benefit from Bridge International Academies Affordable Quality Education." *BusinessDay*, February 1, 2017.

Rainey, Lydia. Making Sense of Charter School Studies, CRPE Reinventing Public Education, 2015. Retrieved from http://www.crpe.org/sites/default/files/mediaguide-charterschoolstudies-apr2015_0.pdf.

"Universal Basic Education in Nigeria." Center for Public Impact.
https://www.centreforpublicimpact.org/case-study/universal-basic-education-nigeria/.

Zhukovskii, Mikhail. "Population Distribution by Age 2006." Nigeria Data Portal.
http://nigeria.opendataforafrica.org/xlomyad/population-distribution-by-age-2006?state=Edo.

Appendix

Appendix A. Study Details

A1. Sample Selection

Schools and replacement schools were selected and matched based on the data available from the Edo State School Census in 2017, focusing on variables that exhibited variance across schools.

Step 1. Identify 30 EdoBEST Schools Representative of Edo

From the 262 EdoBEST schools, we selected a sample of 30 schools that was representative of all 1,111 Edo state primary schools. This entailed first randomly selecting 400 samples of 30 EdoBEST schools each, ensuring a geographical spread across LGAs. For each of the samples, we conducted t-tests across the 13 matching characteristics; any samples that had more than one characteristic different was omitted. We then randomly selected one of the three remaining samples for our list of 30 representative treatment schools.

Step 2. Identify 30 Matching Control Schools

With our list of 30 EdoBEST Schools, we utilized a multivariate-distance matching estimate[27] to identify the "closest" comparison school within the same LGA. After identifying the matched pairs, we used a propensity score estimator to identify a third school that could act as a control both for the EdoBEST school and its matched pair. In other words, we created two sets of control schools that matched both the EdoBEST treatment sample and each other. From each pair of matching control schools, we randomly selected one to serve as the control for the study.[28]

Step 3. Identify Replacement Schools

It was imperative that our EdoBEST and status quo schools had P3 and P4 pupils; however, enrollment may have shifted since the census was conducted. Therefore, we identified a list of treatment and control replacement schools using a propensity score estimate. A total of twelve schools were swapped with replacement schools over the course of study implementation.[29]

The final sample of treatment versus control schools are balanced on pre-intervention characteristics and resemble schools across Edo state. See *Figure 30*.

---

27 The multivariate-distance matching estimate uses a Mahalanobis distance. This is equivalent to a Euclidean distance measurement on data that has been transformed through Principal Component Analysis. In other words, it adjusts for the fact that characteristics are correlated and scales the characteristics to become standardized.

28 Note that schools with specific characteristics (such as schools for the mentally handicapped or for the blind) were disqualified from selection as a control.

29 One EdoBEST school in Akoko Edo LGA was replaced with a school in Esan West LGA.

Figure 30. Pre-Intervention School Characteristics of Study Sample, All EdoBEST Schools, and All Edo Schools

Characteristic	Statistic	EdoBEST "Treatment" Sample	Status Quo "Control" Sample	All EdoBEST Schools	All Edo Schools
Rural	Mean	70.0%	76.7%	47.7%	77.0%
	SD	(0.47)	(0.43)	(0.5)	(0.42)
Water Access	Mean	10.0%	10.0%	23.3%	16.0%
	SD	(0.31)	(0.31)	(0.42)	(0.37)
Power Access Experiences Vandalism	Mean	3.3%	3.3%	10.3%	10.4%
	SD	(0.18)	(0.18)	(0.3)	(0.3)
Health Facility Access	Mean	33.3%	26.7%	32.4%	21.5%
	SD	(0.48)	(0.45)	(0.47)	(0.41)
Has Toilets	Mean	30.0%	46.7%	54.6%	47.5%
	SD	(0.47)	(0.51)	(0.5)	(0.5)
School Experiences Vandalism Pct Classes Good	Mean	46.7%	26.7%	58.1%	42.3%
	SD	(0.51)	(0.45)	(0.49)	(0.49)
Commute Less Than 20	Mean	40.0%	30.0%	29.4%	38.7%
	SD	(0.5)	(0.47)	(0.46)	(0.49)
PTR	Mean	21.5	21.6	23.2	21.3
	SD	(11.86)	(10.32)	(9.96)	(12.53)
% Female Teachers Has Water	Mean	74.1%	70.2%	83.7%	67.6%
	SD	(0.2)	(0.26)	(0.17)	(0.28)
% Teachers w/ Advanced Degree	Mean	10.7%	10.9%	15.4%	11.1%
	SD	(0.11)	(0.12)	(0.14)	(0.14)
Classrooms With Blackboards	Mean	52.7	46.2	69.9	45.9
	SD	(40.14)	(29.56)	(39.35)	(36.4)
% Classes in Good Condition Has Power	Mean	42.5%	30.7%	41.4%	32.2%
	SD	(0.43)	(0.34)	(0.41)	(0.39)
% Classrooms with Blackboard	Mean	77.5%	80.9%	81.5%	77.7%
	SD	(0.32)	(0.33)	(0.31)	(0.35)
Count	#	30	30	262	1,111

Note: There are no statistically significant differences between the sampled EdoBEST schools and their control counterparts.

A2. Measurements

Measurements were designed to capture changes along the EdoBEST theory-of-change, in inputs, drivers, and outcomes.

Figure 31. Study Measurement Tools Overview

Instrument	Purpose
Headmaster Interviews	Capture school characteristics, including the number of teachers, pupils, and classrooms for P3 and P4. Administered to each headmaster.
School Observations	Observe when school commenced and ended, as well as whether the P3 and P4 teacher participating in the study was present in the afternoon.
Classroom Observations	Observe pupil and teacher attendance and behavior in the classroom, along with the presence of learning supports.
Teacher Interview	Capture pupil enrollment and self-reports of training and ongoing support.
Maths and Literacy Assessments	Capture P3 and P4 pupil performance on math and English literacy.

a. Inputs

Instruments were designed to capture how the EdoBEST program affected three types of inputs: (1) access to resources, (2) increased capacity building, and (3) ongoing quality assurance.

1. Resources: EdoBEST teachers are provided a teacher computer with lesson guides and headmasters a smartphone. We measure whether the teacher was using an electronic teaching aid during our classroom observation. Second, the EdoBEST program provides more on-level texts for pupils. During the classroom observation, assessors count how many textbooks are visible in the classroom to calculate the pupil-to-textbook ratio.[30] Third, official instructional time was increased by standardizing a daily timetable that scheduled more teaching activities during the school day. To determine whether this input is changed, assessors noted which teachers were still at school at 1:45 and at 2:30 in the afternoon.

2. Capacity Building: EdoBEST teachers undergo a three-week long training and headmasters are trained to provide ongoing coaching to EdoBEST teachers. During teacher interviews, we asked teachers how many times they attended training in the past year. We also asked if they were observed by the headmaster during the past week and whether the headmaster provided feedback after the observation. Finally, leadership and development officers are to visit schools approximately every 1-2 weeks to provide coaching and feedback. We therefore also ask if someone from outside the school has visited and provided feedback.

---

30 We chose to count visible textbooks rather than textbooks in use because some lessons in both EdoBEST and status quo schools do not require the use of textbooks. Assessors were instructed to scan the classroom for textbooks visible on pupils' or teachers' desks or on bookshelves. Textbooks locked in cabinets not visible to the assessors were not counted.

3. Quality Assurance: Quality assurance officers and academic field officers visit EdoBEST schools on a rotating basis. This mechanism aims to set quality standards for EdoBEST schools, provide feedback to improve instructional quality, and provide information to the central support office. During teacher interviews, we ask teachers whether they have been visited in the past month by someone outside of the school and, if so, whether they received feedback during that visit.

b. Drivers

Under EdoBEST's theory of change, learning is impacted by drivers under three main categories: (1) increased time to learn, (2) enhanced academic environment, and (3) improved instructional quality.

1. Increased Time to Learn: There are many mechanisms through which pupils have more time to learn in the EdoBEST classroom. First, accountability mechanisms within the EdoBEST system (regular visits by quality assurance, check-ins with the technology team) and the time saved with pre-written lesson guides will positively impact teacher attendance. Our unannounced classroom observations began at 9 AM. After a ten-minute buffer period, we measured whether teachers were present during each classroom observation period. Second, the EdoBEST program may have positive impacts on pupil attendance as lessons are designed to align with pupil levels. We measure pupil attendance by looking at the percent of pupils counted in class during the classroom observation divided by the number of enrolled pupils as reported in the teacher interview. Finally, changes in classroom management procedures should result in increased time on task; we estimate the number of minutes pupils were focused during each observation period.

2. Improved Environment: The EdoBEST program aims to improve the environment of the classroom by reducing misbehavior and corporal punishment, engaging parents, and increasing praise within the classroom. We measured the number of minutes lost due to misbehavior during the classroom observation. Assessors also recorded whether the teacher scolded, belittled, or used corporal punishment on pupils at any point during the classroom observation. Finally, assessors counted how often teachers used generic and specific (narrating the postive) praise during the classroom observation.

3. Improved Instructional Quality: The EdoBEST program uses pedagogical techniques proven to result in more learning. When implemented correctly, these techniques result in more engaged pupils that eagerly follow directions, greater participation in the lesson, and continuous feedback to pupils. Assessors estimated the number of pupils engaged in the lesson, the number of pupils working hard during the lesson, the number of times pupils were called on and, if called on, whether they heard if their response was correct. Assessors also marked whether the lesson included independent practice time.

c. Outcomes

The exam includes both revision and at-level questions. Data was collected at the item-level, which allowed us to create sub-totals based on item difficulty. Exams were piloted and item analysis was used to ensure exams included items of varying difficulty levels and discriminatory power. A break-down on the type of questions under each category can be found below.

Figure 32. Assessment Question Type Details

Subject	Sub-score	Content
Maths	Revision	Counting (1), quantity discrimination (1), inequalities (2), addition (5), subtraction (4)
	At-level	Reading time on a clock (3), days in a month (2), time facts (6), lines of symmetry (2), 3D shapes (2), estimation (2), reading data on pictograms (6), computation with units of weight (5)
	Total	Sum of revision and at-level content
Literacy	Revision	*Listening:* letter identification (3), letter sounds (3), word identification (6), initial sounds (3), final sounds (1), *Reading:* word identification (3), reading comprehension (3)
	At-level	Phonics (6), reading comprehension (8)
	Total	Sum of revision and at-level phonics and reading comprehension

A3. Training and Data Collection

From July 11th through 13th we trained assessors on administering headmaster/teacher interviews, conducting school and classroom observations, proctoring numeracy and literacy exams, and inputting data. To ensure reliable classroom observation data, assessors were trained on a simple rubric that recorded information on pupil engagement, participation, diligence and the teacher's classroom management. After training assessors with recorded classroom observations, assessors paired up and conducted the classroom observation in EdoBEST schools not in the study. At the end of the training, assessors reached an inter-rater reliability estimate of 0.74.

Data was collected from July 16th to July 23rd. Assessors visited a school over a two-day period.31 On day one, assessors conducted the headmaster interview, observed the P3 classroom, conducted the P3 teacher interview, and proctored the literacy exam for P3 and P4 pupils. In the afternoon, assessors graded the literacy exam and inputted interview and observation data. On day two, assessors observed the P4 classroom, conducted the P4 teacher interview, and proctored the maths exam for P3 and P4 pupils.32 In the afternoon, assessors inputted item-level data for the both the literacy and the numeracy exam.

Frequently, Edo schools include multiple classrooms within the same grade. As such, we required a sampling methodology to select a class randomly from the school. We trained assessors to identify the full list of P3 and P4 teachers at the school. From the list of teachers, the first teacher alphabetically was chosen. Then, assessors would ask the headmaster if the teacher was present. If yes, the assessor would ask to be pointed to the classroom. If no, the assessor would move on to the next teacher alphabetically.

Assessors were provided with enough assessments to proctor the exam for all pupils in the classroom. However, if there was a mismatch between the number of assessments provided and the number of pupils in the classroom, assessors were trained in a simple randomization strategy to select a sub-sample of pupils. This was only required once during the study.

31 Generally, we assigned assessors to visit a pair of schools (EdoBEST and control schools) when possible. However, due to schedule constraints, six schools were split across multiple assessors.
32 The multiple-choice section of the literacy exam and the entire mathematics exam were recorded using a bubble sheet. This was the first time the majority of pupils in EdoBEST and control schools used a bubble sheet. Each assessor was trained on teaching the pupils how to fill out the bubble sheet and, with the help of the homeroom teacher, supported pupils in filling out the bubble sheet.

We utilized electronic survey collection to monitor data collection and address data inconsistencies as they arose. In the classroom observation survey, for example, assessors recorded the number of pupils called on and (out of those called on) how many pupils heard whether their answer was correct or incorrect. If the second answer was larger than the first, we asked assessors to double check their data entry and fix any errors. By monitoring data collection in real-time, we were able to address inconsistencies before they were buried under multiple days of work.

A4. Limitations Explained

Power: The study will be powered to detect a roughly 3/8 standard deviation difference. As a rule of thumb, studies in education are usually powered to detect a 1/5 standard deviation difference for a full year intervention. This means that we will only have statistically significant findings if EdoBEST drives almost twice the learning in a third of the time. As such, it is important to remember up front that null learning differences does *not* necessarily mean no learning; rather, learning gains need more time to accrue. For that reason, we will measure both learning gains and assess differences in the conditions that lead to academic performance. We believe that these conditions, or inputs to support learning, are likely to change more quickly and as such will be easier to detect than changes in pupil performance.

Selection Bias: Schools were chosen to be a part of the EdoBEST pilot and this may result in measuring schools that were "most likely to succeed", even without the intervention. We tried to mitigate this threat in two ways. First, we select control schools that most closely matched the observable characteristics to the treatment sample. Second, we include P4 pupils in the study to check for systematic differences between pupils in our treatment and control schools. This sampling strategy is described above in *Appendix A1. Sample* .

Generalizability: Because schools were chosen to be a part of the EdoBEST pilot, they may not be representative of all schools in Edo state. Instead of choosing a treatment sample representative of the EdoBEST pilot schools, we selected a sample of schools that are most representative of Edo State more generally. This will improve our ability to generalize the results of our study beyond the pilot.

Spillover Effects: The EdoBEST pilot extended from Primary 1 up to Primary 3. We exploit partial treatment of schools to isolate the impact. We acknowledge, however, that changes in teaching methods and managerial processes could impact other classrooms within a school, potentially positively biasing performance of P4 EdoBEST classrooms. It is also possible that parents may be eager for their pupils to participate in the EdoBEST classrooms which could impact the makeup of both the P3 and P4 classrooms in EdoBEST schools. Because we do not have data on pupil placement and teacher performance before the intervention, we are unable to analyze how these spillover effects bias our estimates of the EdoBEST impact. We theorize, though, that the EdoBEST program would have a positive impact on P4 pupils, which would result in an underestimation of our impact estimate.

Appendix B. Regression Output

Across regression outputs, the highlighted EdoBEST effect is the estimate of the impact of the EdoBEST program on the outcome variable. When the outcome variable is a continuous variable, such as observed length of school, the result should be interpreted as the marginal effect in the outcome unit. When the outcome variable is a percent, such % of teachers attended training, the result should be interpreted as a percentage point increase.

B1. Inputs to Support Learning

Table 1. Inputs: Two-way Fixed Effects w/ Controls

	Teacher Support					Learning Materials	
	(1) % of Teachers Attended Training	(2) # of Trainings Attended	(3) % of Teachers Received Outsider Feedback	(4) % of Teachers Received Headmaster Feedback	(5) % of Teachers Used Electronic Device	(6) % of Classrooms w/ Textbooks Visible	(7) # of Textbooks Accessible to Pupils
EdoBEST Effect	0.73**	0.76**	0.25	-0.07	1.00**	0.33+	8.98*
	(0.11)	(0.18)	(0.18)	(0.18)	(0.00)	(0.17)	(4.37)
EdoBEST Participating School	0.13	0.14	0.09	0.00	-0.00**	0.00	-1.59
	(0.08)	(0.11)	(0.14)	(0.14)	(0.00)	(0.13)	(2.63)
P3	0.08	0.11	0.03	0.25+	-0.00**	-0.04	2.09
	(0.07)	(0.12)	(0.13)	(0.12)	(0.00)	(0.13)	(1.85)
Class Size (10s of students)	0.00	0.00+	0.00	0.00	0.00	0.00	0.11
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)
# of Other Teachers Present	-0.11*	-0.17**	-0.03	-0.01	0.00	0.03	5.77+
	(0.04)	(0.06)	(0.10)	(0.09)	(0.00)	(0.09)	(3.38)
Constant	-0.01	-0.05	0.43**	0.47**	0.00**	0.56**	-2.42
	(0.05)	(0.08)	(0.12)	(0.13)	(0.00)	(0.13)	(4.05)
Number of Observations	111	111	111	111	111	119	119
R-squared	0.73	0.59	0.27	0.29	1.00	0.25	0.38

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors are clustered at the school level and shown in parentheses. All regressions include LGA controls.

B2. Drivers of Learning

Table 2. More Time: Two-way Fixed Effects w/ Controls

	Teachers & Pupils Present						Instructional Time	
	(1) Hours in the School Day[a]	(2) % of Teachers Present in Morning	(3) % of Teachers Present at 1:45 PM	(4) % of Teachers Present at 2:30 PM	(5) % of Pupils in Attendance	(6) % of Classrooms with Either No Teacher or Pupils Present	(7) 'Time Most Pupils Focused' Rating (-1 to 1)	(8) 'Time Lost to Misbehavior Rating (-1 to 1)
EdoBEST Effect	0.75**	0.09	0.19	0.38**	0.08	-0.13	0.33+	-0.05
	(0.16)	(0.08)	(0.13)	(0.13)	(0.07)	(0.08)	(0.19)	(0.13)
EdoBEST Participating School		-0.05	0.04	0.39**	0.02	0.05	-0.02	0.09
		(0.08)	(0.10)	(0.11)	(0.05)	(0.08)	(0.15)	(0.12)
P3		0.00	-0.03	0.02	-0.03	0.03	0.06	-0.10
		(0.05)	(0.10)	(0.06)	(0.05)	(0.05)	(0.14)	(0.09)
Class Size (10s of students)	0.01	-0.00	0.00	0.00	-0.00*	0.00	-0.01**	-0.00
	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Other Teachers Present	0.03	0.06	0.09+	0.00	0.04	-0.06	0.22**	0.03
	(0.10)	(0.04)	(0.05)	(0.06)	(0.03)	(0.04)	(0.08)	(0.06)
Constant	5.51**	0.96**	0.71**	0.01	0.80**	0.05	0.69**	-0.83**
	(0.18)	(0.05)	(0.09)	(0.07)	(0.05)	(0.04)	(0.13)	(0.09)
Number of Observations	60	119	119	119	119	119	357	330
R-squared	0.55	0.24	0.31	0.63	0.24	0.31	0.24	0.10

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors are clustered at the school level and shown in parentheses. All regressions include LGA controls.
a) The intervention here happened at the school level, and as such a full difference-in-difference estimate cannot be calculated. Class size here is calculated as total students and the number of teachers present is the total P3 and P4 teachers at the school.

Table 3. Better Environment: Two-way Fixed Effects w/ Controls

	Punishment & Negative Culture		Praise & Positive Discipline							
	(1) % of Teachers Who Scolded or Belittled	(2) % of Teachers who Used Corporal Punishment	(3) % of Teachers Who Praised Pupils	(4) % of Teachers Who Praised at Least 1-2 Pupils	(5) % of Teachers Who Praised at Least 3 Pupils	(6) % of Teachers Who Praised the Whole Class Together	(7) % of Teachers Who Positively Narrated the Behavior of Pupils	(8) % of Teachers Who Narrated the Positive for 1-2 Pupils	(9) % of Teachers Who Narrated the Positive for 3 or More Pupils	(10) % of Teachers Who Narrated the Positive to the Whole Class Together
EdoBEST Effect	0.06	0.07	0.54**	0.33*	0.34*	0.65**	0.35**	0.26*	0.18*	0.28**
	(0.16)	(0.12)	(0.16)	(0.16)	(0.14)	(0.13)	(0.12)	(0.11)	(0.07)	(0.09)
EdoBEST Participating School	-0.13	-0.18*	-0.03	-0.11	-0.14	0.09	0.07	0.08	0.03	0.02
	(0.13)	(0.08)	(0.13)	(0.13)	(0.10)	(0.09)	(0.06)	(0.06)	(0.03)	(0.03)
P3	-0.07	-0.04	-0.04	-0.02	-0.02	0.04	0.06	0.07	-0.01	-0.01
	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)	(0.07)	(0.06)	(0.06)	(0.03)	(0.03)
Class Size (10s of students)	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Other Teachers Present	-0.15**	-0.04	-0.02	0.08	0.22**	-0.06	-0.01	-0.04	-0.08*	-0.02
	(0.05)	(0.04)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.03)	(0.04)
Constant	0.35**	0.21*	0.45**	0.50**	0.24**	0.02	0.03	0.03	0.01	-0.02
	(0.12)	(0.08)	(0.10)	(0.10)	(0.09)	(0.10)	(0.06)	(0.05)	(0.04)	(0.05)
Number of Observations	111	111	111	111	111	111	111	111	111	111
R-squared	0.38	0.31	0.49	0.45	0.52	0.58	0.41	0.37	0.30	0.37

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors are clustered at the school level and shown in parentheses. All regressions include LGA controls.

Table 4. Instructional Quality: Two-way Fixed Effects w/ Controls

	Pupils Follow Directions	Pupils Work Hard Independently		Pupils Answer Questions & Get Feedback		
	(1)	(2)	(3)	(4)	(5)	(6)
	'Pupils Follow Directions' Rating (-1 to 1)	% Obs. Where Independent Practice Occurred	'Pupils Work Hard During Independent Practice' Rating (-1 to 1)	% Pupils Called On in Front of Whole Class	% of Pupils Who Received Concrete Feedback in Front of Whole Class	% Pupils Who Have Work Checked During Independent Practice
EdoBEST Effect	0.27+	0.11	0.56*	0.07	0.06	0.27**
	(0.14)	(0.14)	(0.23)	(0.07)	(0.05)	(0.07)
EdoBEST Participating School	-0.10	0.08	-0.01	-0.04	-0.03	-0.01
	(0.11)	(0.10)	(0.20)	(0.05)	(0.04)	(0.07)
P3	-0.00	0.15	-0.27	-0.02	-0.01	-0.08
	(0.12)	(0.09)	(0.20)	(0.04)	(0.03)	(0.05)
Class Size (10s of students)	-0.00	-0.00	-0.00	-0.00**	-0.00**	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Other Teachers Present	0.07	-0.02	-0.04	0.06*	0.04+	-0.02
	(0.06)	(0.07)	(0.13)	(0.03)	(0.02)	(0.05)
Constant	0.80**	0.32**	0.82**	0.29**	0.19**	0.28**
	(0.10)	(0.07)	(0.21)	(0.04)	(0.03)	(0.07)
Number of Observations	327	330	139	230	230	139
R-squared	0.10	0.16	0.23	0.22	0.25	0.41

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors are clustered at the school level and shown in parentheses. All regressions include LGA controls.

B3. Learning Achievement

Table 5. Maths Raw (% Questions Correct)

	Reduced Form			Fixed Effects			Fixed Effects with Controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Revision	At-Level	Total	Revision	At-Level	Total	Revision	At-Level	Total
EdoBEST Effect on % Math Scores	0.06	0.00	0.02	0.07*	0.02	0.04+	0.07*	0.02	0.04+
	(0.04)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
EdoBEST Participating School				-0.01	-0.01	-0.01	-0.02	-0.02	-0.02
				(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Primary 3				-0.10**	-0.05**	-0.07**	-0.10**	-0.05**	-0.07**
				(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Female							0.03**	0.01	0.02*
							(0.01)	(0.01)	(0.01)
Age							-0.00	-0.00	-0.00
							(0.00)	(0.00)	(0.00)
Class Size (10s of students)							-0.01	-0.01*	-0.01+
							(0.01)	(0.00)	(0.00)
Number of Other Teachers							0.02	0.03**	0.02*
							(0.02)	(0.01)	(0.01)
Constant	0.58**	0.26**	0.38**	0.68**	0.31**	0.44**	0.74**	0.34**	0.48**
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.05)	(0.02)	(0.03)
Number of Observations	1,640	1,640	1,640	3,162	3,162	3,162	3,162	3,162	3,162
R-squared	0.01	0.00	0.01	0.02	0.02	0.03	0.10	0.10	0.11

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors clustered at the school level are shown in parentheses. LGA controls are included for regressions (7), (8), and (9).

Table 6. Maths STD

	Reduced Form			Fixed Effects			Fixed Effects with Controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Revision	At-Level	Total	Revision	At-Level	Total	Revision	At-Level	Total
EdoBEST Effect on Standardized Math Scores	0.21	0.03	0.15	0.27*	0.13	0.24+	0.27*	0.12	0.24+
	(0.14)	(0.12)	(0.13)	(0.12)	(0.15)	(0.14)	(0.12)	(0.14)	(0.13)
EdoBEST Participating School				-0.05	-0.10	-0.09	-0.06	-0.18	-0.13
				(0.13)	(0.15)	(0.15)	(0.09)	(0.11)	(0.11)
Primary 3				-0.38**	-0.37**	-0.44**	-0.38**	-0.35**	-0.43**
				(0.08)	(0.09)	(0.09)	(0.07)	(0.09)	(0.07)
Female							0.13**	0.05	0.11*
							(0.04)	(0.04)	(0.04)
Age							-0.02	-0.01	-0.02
							(0.01)	(0.01)	(0.01)
Class Size (10s of students)							-0.03	-0.06*	-0.05+
							(0.03)	(0.03)	(0.03)
Number of Other Teachers							0.06	0.20**	0.15*
							(0.06)	(0.06)	(0.06)
Constant	-0.18*	-0.18*	-0.21*	0.20*	0.19	0.23*	0.41*	0.40*	0.48*
	(0.09)	(0.08)	(0.09)	(0.08)	(0.12)	(0.11)	(0.19)	(0.19)	(0.19)
Number of Observations	1,640	1,640	1,640	3,162	3,162	3,162	3,162	3,162	3,162
R-squared	0.01	0.00	0.01	0.02	0.02	0.03	0.10	0.10	0.11

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors clustered at the school level are shown in parentheses. LGA controls are included for regressions (7), (8), and (9).

Table 7. Literacy Raw (% Questions Correct)

	Reduced Form			Fixed Effects			Fixed Effects with Controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Revision	At-Level	Total	Revision	At-Level	Total	Revision	At-Level	Total
EdoBEST Effect on % English Literacy Scores	0.06*	0.03	0.05+	0.07*	0.02	0.05+	0.06*	0.02	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)
EdoBEST Participating School				-0.01	0.01	-0.00	-0.02	-0.01	-0.02
				(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Primary 3				-0.10**	-0.09**	-0.10**	-0.11**	-0.10**	-0.11**
				(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Female							0.03**	0.02**	0.03**
							(0.01)	(0.01)	(0.01)
Age							-0.01**	-0.01*	-0.01**
							(0.00)	(0.00)	(0.00)
Class Size (10s of students)							0.01+	0.01	0.01
							(0.01)	(0.01)	(0.01)
Number of Other Teachers							0.03+	0.00	0.02
							(0.02)	(0.02)	(0.02)
Constant	0.53**	0.34**	0.47**	0.62**	0.43**	0.57**	0.66**	0.47**	0.60**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)	(0.05)	(0.04)
Number of Observations	1,551	1,551	1,551	3,031	3,031	3,031	3,031	3,031	3,031
R-squared	0.03	0.00	0.02	0.04	0.03	0.04	0.22	0.17	0.23

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors clustered at the school level are shown in parentheses. LGA controls are included for regressions (7), (8), and (9).

Table 8. Literacy STD

	Reduced Form			Fixed Effects			Fixed Effects with Controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Revision	At-Level	Total	Revision	At-Level	Total	Revision	At-Level	Total
EdoBEST Effect on Standardized Literacy Scores	0.33*	0.13	0.28+	0.36*	0.10	0.28+	0.33*	0.08	0.25
	(0.15)	(0.14)	(0.16)	(0.14)	(0.16)	(0.14)	(0.16)	(0.17)	(0.17)
EdoBEST Participating School				-0.03	0.03	-0.01	-0.11	-0.03	-0.09
				(0.20)	(0.20)	(0.22)	(0.14)	(0.15)	(0.16)
Primary 3				-0.52**	-0.40**	-0.53**	-0.61**	-0.45**	-0.61**
				(0.10)	(0.06)	(0.08)	(0.09)	(0.08)	(0.08)
Female							0.17**	0.11**	0.16**
							(0.04)	(0.04)	(0.04)
Age							-0.05**	-0.03*	-0.05**
							(0.01)	(0.01)	(0.01)
Class Size (10s of students)							0.08+	0.03	0.07
							(0.04)	(0.04)	(0.05)
Number of Other Teachers							0.16+	0.01	0.11
							(0.09)	(0.11)	(0.10)
Constant	-0.25**	-0.19+	-0.26**	0.27*	0.21+	0.27*	0.45+	0.37	0.46+
	(0.09)	(0.10)	(0.09)	(0.13)	(0.11)	(0.13)	(0.23)	(0.22)	(0.24)
Number of Observations	1,551	1,551	1,551	3,031	3,031	3,031	3,031	3,031	3,031
R-squared	0.03	0.00	0.02	0.04	0.03	0.04	0.22	0.17	0.23

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors clustered at the school level are shown in parentheses. LGA controls are included for regressions (7), (8), and (9).

B4. Heterogeneous Impact

Table 9. Raw Assessment Results Parsed by Gender

	Maths			English Literacy		
	(1)	(2)	(3)	(4)	(5)	(6)
Revision	Revision	At-Level	Total	Revision	At Level	Total
Female*EdoBEST*P3	0.07**	0.02*	0.04**	0.05**	0.07**	0.06**
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Female*EdoBEST	-0.00	-0.01	-0.01	0.02	0.00	0.02
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Female*P3	0.02	0.01	0.01	0.01	-0.00	0.01
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Female	0.04**	0.01	0.02**	0.04**	0.02	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
EdoBEST*P3	0.03	0.00	0.01	0.04	-0.02	0.01
	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)
EdoBEST	0.00	-0.01	-0.01	-0.01	0.00	-0.01
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
P3	-0.10**	-0.05**	-0.06**	-0.10**	-0.09**	-0.10**
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Number of Other Teachers	0.02	0.03**	0.02*	0.03+	0.00	0.02
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Age	-0.01	-0.00	-0.00	-0.01**	-0.01*	-0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Class Size (10s of students)	-0.01	-0.01*	-0.01+	0.01+	0.01	0.01
	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Constant	0.74**	0.34**	0.48**	0.65**	0.47**	0.60**
	(0.05)	(0.03)	(0.03)	(0.04)	(0.05)	(0.04)
Number of Observations	3,162	3,162	3,162	3,031	3,031	3,031
R-squared	0.10	0.10	0.11	0.22	0.18	0.23

Note: Two asterisks (**) denote statistical significance at the 99% confidence level, one asterisk (*) at the 95% confidence level, and a cross (†) at the 90% confidence level. Robust standard errors clustered at the school level are shown in parentheses. All regressions include controls for LGAs.