Oxford Policy Management

ESSPIN Composite Survey 2

Kaduna State report

Stuart Cameron and Nicola Ruddle

June 2015

Acknowledgements

The authors are very grateful to a number of people who helped in producing this note. Ravi Somani, Sourovi De and Peter-Sam Hill worked on parts of the data analysis. David Megill reviewed weights from the first Composite Survey and produced the sampling weights application for the second Composite Survey. Allan Findlay provided support with school identification and the annual school census. From ESSPIN, Jake Ross, Kayode Sanni, Fatima Aboki, Sandra Graham, Pius Elumeze, Lilian Breakell, John Kay and Simon Thomson reviewed preliminary notes and drafts and provided valuable comments. Gratitude is also due to everyone who conducted the second Composite Survey, including the staff of Oxford Policy Management and ESSPIN, state coordinators, and the staff of the State Universal Basic Education Boards, who did the data collection; and not least to the large number of headteachers, teachers and students who took the time to participate in the study.

This survey exercise is being carried out by Oxford Policy Management for the Education Sector Support Programme in Nigeria (ESSPIN), which is funded by the UK Department for International Development and managed by a consortium led by Cambridge Education. The survey project manager is Stuart Cameron. For further information contact stuart.cameron@opml.co.uk.

Oxford Policy Management Limited
Registered in England: 3122495

6 St Aldates Courtyard
38 St Aldates
Oxford OX1 1BN
United Kingdom

Tel +44 (0) 1865 207 300
Fax +44 (0) 1865 207 301
Email admin@opml.co.uk
Website www.opml.co.uk

Executive summary

This report presents findings for Kaduna State from the first and second rounds of the Education Sector Support Programme in Nigeria (ESSPIN) Composite Survey (CS1 and CS2), conducted in 2012 and 2014, respectively. The survey covered a wide range of indicators at the teacher, headteacher, school-based management committee (SBMC), and pupil levels, in an attempt to understand how schools in ESSPIN states are changing over time and whether schools which receive ESSPIN interventions are working better than those which do not.

The ESSPIN model involves training for teachers, headteachers, and community members on SBMCs. In Kaduna a pilot phase of ESSPIN began in 2009/10, and in 2012/13 and 2013/14 the interventions were rolled out to two more groups of schools. This expansion involved a changed model for delivering training, which made the scale-up possible and located training closer to or within the schools. By the time of the 2014 survey, 27% of schools in Kaduna had received at least one year of the full package of ESSPIN support. Schools which benefited from ESSPIN only in 2013/14 are treated as non-ESSPIN for the purpose of this report, since the support is not expected to have had an effect yet.

There have been large increases in pupil enrolment in Kaduna since 2009, and there has also been a rise in violent conflict in the state since 2010. Both of these factors may have affected the ability of schools to deliver good learning outcomes for enrolled children. Rapid increases in enrolment tend to increase class sizes and make it harder for teachers to ensure that all children are learning, while violent conflict can obviously have severe impacts on school functioning, including deterring children or teachers from attending and making it harder for the community to engage with the school.

The survey data show that there was substantial variation between the results of pupils, teachers, headteachers and schools in Kaduna, creating more uncertainty in the estimation of indicators and, importantly, in the estimation of differences in indicators between groups or over time. As a result, it is difficult to identify confidently differences in the indicators (positive or negative) and therefore to establish to what extent the ESSPIN programme has had the expected impact. The main findings are as follows:

Teacher competence: Teachers in ESSPIN schools appear to perform better than in non-ESSPIN schools. Two behaviours – the use of teaching aids and encouraging children by giving more praise than reprimands – are more likely in particular to be found among teachers in ESSPIN schools than in others. For Kaduna as a whole, the level of teacher competence has fallen, with the proportion of teachers reaching our competence standard decreasing from 76% in 2012 to 62% in 2014. Teachers correctly answered only around one in five grade 5 English questions, and one in three grade 5 maths questions. Looking at the change over time, we find no evidence that teachers who were trained under ESSPIN improved faster (or worsened less) than teachers in non-ESSPIN schools between the two surveys.

Headteacher effectiveness: The proportion of headteachers meeting our standard for effectiveness rose from 8.5% in 2012 to 13% in 2014, although this change was not statistically significant and disappeared when we used an alternative indicator (number of effectiveness criteria met). There was little difference between the results in ESSPIN schools and in non-ESSPIN schools in terms of our overall effectiveness measure, but headteachers in ESSPIN schools were much more likely to have carried out lesson observations and to have had professional development meetings with their teachers. Even in ESSPIN schools, only 10% of headteachers could show our interviewers lesson observation sheets and only 20% had a record of professional development meetings, but in non-ESSPIN schools almost none had these. The pace of change in

headteacher effectiveness between 2012 and 2014 was about the same in ESSPIN and non-ESSPIN schools.

School development planning: ESSPIN schools were better at school development planning than control schools, although they still met only 1.3 out of five criteria on average. Around 40% of ESSPIN schools had carried out a self-evaluation process and had a school development plan (SDP) of some form, compared to only 4% of non-ESSPIN schools. The average level of school development planning has not changed significantly in the two years to 2014 in either ESSPIN or non-ESSPIN schools.

School inclusiveness: ESSPIN schools appear to perform better than non-ESSPIN schools on a number of indicators relating to inclusiveness (although the differences were for the most part, not statistically significant). In Kaduna as a whole, however, schools appear to have become less inclusive between 2012 and 2014, with only 11% now reaching our inclusiveness standard, compared to 23% in 2012. Similar declines in inclusiveness indicators are found in both ESSPIN and non-ESSPIN schools, but ESSPIN schools remain more inclusive.

SBMC functionality and inclusiveness: SBMCs are found in nearly all schools in Kaduna, but in non-ESSPIN schools do not appear to be functioning well. In ESSPIN schools, 28% of SBMCs could show a record of having met more than once in the past year, compared to only 1.5% of SBMCs in non-ESSPIN schools. Our overall standard for a functioning SBMC was met by 35% of ESSPIN schools and 18% of non-ESSPIN schools. SBMCs in ESSPIN schools were also significantly more likely to be inclusive of women (14%, compared to 2% in non-ESSPIN schools), and of children (7%, compared to none of the non-ESSPIN schools). There was little change over time in either group of schools on any of these indicators, except that many more SBMCs in 2014 had taken action for commonly excluded groups (11% in 2014, compared to 2% in 2012) or raised the issue of children's exclusion (17% vs. 4%). The extent to which SBMCs encouraged children's participation had also increased, although it remained very low: 1% of schools met our standard for inclusiveness of children in SBMC activities.

School quality: We measure overall quality by combining some of the above indicators – headteacher effectiveness, teacher competence, school development planning, and SBMC functionality – into a single indicator: 15% of ESSPIN schools met our overall standard for school quality, while none of the control schools did. Comparison of the results over time suggests that there was an improvement in quality in ESSPIN schools between 2012 and 2014, while other schools in the state may have become worse, although these differences do not reach statistical significance. For the state as a whole, there has been little change over time in school quality.

Pupil learning: Tests were conducted in literacy and numeracy for pupils in grades 2 and 4. In Kaduna the average test scores fell for all four tests between 2012 and 2014. The average test scores range from 21% in grade 4 literacy to 34% in grade 2 numeracy. Test scores in ESSPIN schools appear to be slightly higher than in other schools, but the difference is not statistically significant for any of the tests. The declines in test scores over time were of a similar magnitude in both ESSPIN and other schools.

Overall, ESSPIN schools are functioning significantly better than other schools in the state in terms of most of our indicators, although the differences do not always reach statistical significance. But taking the results for Kaduna as a whole, there have been slight improvements in SBMC inclusiveness and children's participation in SBMCs; no significant change in headteacher effectiveness, SBMC functionality, or overall school quality; and a worsening in teacher competence, school inclusiveness, and pupil test results. These two findings – the better quality of ESSPIN schools compared to control schools, and the deterioration over time in the state as a

whole – can be reconciled by noting that ESSPIN interventions have been rolled out to around one-quarter of schools in the state, so the remaining three-quarters have not yet benefited.

Although ESSPIN schools appear to be more functional than control schools in many respects, these changes have not yet translated into better pupil test results. A possible explanation for this is that teachers' limited subject knowledge acts as a bottleneck, preventing learning outcomes from improving even in schools that work relatively well in other ways. Kaduna's results also come in a context of underlying conflict and violence in the state, an extraordinary situation for communities, teachers and pupils to be living and working in. The impact on the ability to teach, learn, and provide functioning, effective leadership and management cannot be quantified here but is likely to have had considerable influence on the results over the period.

Box 1. The good and bad news from the Composite Surveys in Kaduna

Positive results in this report include:

- ESSPIN schools are functioning better than non-ESSPIN schools in terms of school development planning, inclusiveness of the school as a whole, functionality and inclusiveness of SBMCs, and overall quality.
- Teachers trained by ESSPIN are more competent than teachers in schools which have not benefited from ESSPIN, and on some criteria headteachers are also better in ESSPIN than in non-ESSPIN schools.
- SBMCs in Kaduna as a whole were taking more action on the exclusion of children from school in 2014 than in 2012, and were also more inclusive of children in their own activities.
- The overall quality of ESSPIN schools appears to have improved between 2012 and 2014.

Some challenges identified in the report include:

- Pupil test scores fell between 2012 and 2014, and similar declines were found in both ESSPIN and non-ESSPIN schools. On average, children could answer between 21% and 34% of our literacy and numeracy test questions correctly.
- Taking schools in the state as a whole, teacher competence and school inclusiveness have worsened since 2012.
- Headteacher effectiveness and school development planning have not changed significantly between 2012 and 2014.
- The quality of non-ESSPIN schools in the state appears to have worsened between 2012 and 2014, and the quality of schools in the state as a whole has not changed significantly.

Table of contents

List of tables, figures and boxes

List of abbreviations

ACLED	Armed Conflict Location & Event Data Project
CBO	Community-Based Organisation
CS1	Composite Survey 1
CS2	Composite Survey 2
ESSPIN	Education Sector Support Programme in Nigeria
LGEA	Local Government Education Authority
L2	grade 2 literacy test
L4	grade 4 literacy test
N2	grade 2 numeracy test
N4	grade 4 numeracy test
SBMC	School-based management committee
SDP	School development plan
SIP	School Improvement Programme

1 Introduction

The aims of ESSPIN's Composite Surveys are to assess the effects of ESSPIN's integrated School Improvement Programme (SIP) and report on quality of education in the six ESSPIN-supported states. This report focuses on the key findings for Kaduna State. The surveys address five output indicators: teacher competence, headteacher effectiveness, school development planning, SBMC functionality, and inclusive practices in schools. It also addresses one outcome indicator – school quality – and one impact indicator – pupil learning achievement.

The second round of the Composite Survey (CS2), conducted in 2014, aimed to provide post-intervention data which could be compared to data from the first round of the survey (CS1) collected in 2012, in order to evaluate the extent of improvements in key indicators and gauge programme success. A further survey will be conducted in 2016 to again assess the impact of the interventions.

This report, focusing on Kaduna State, presents findings from CS2 and comparisons between CS1 and CS2, covering all of ESSPIN's output, outcome and impact indicators.

1.1 ESSPIN's SIP

The ESSPIN programme aims to bring about better learning outcomes for children of basic education school age in six states, with a range of activities at the national, state, local and school levels. It has four output streams, focusing on (i) strengthening federal government systems; (ii) increasing the capability of state and local governments for the governance and management of schools; (iii) strengthening capability of primary schools to provide improved learning outcomes; and (iv) improving inclusion policies and practices in basic education (ESSPIN, 2013c).

Under the third of these output streams, ESSPIN's SIP aims to provide and support the use of structured materials that ensure teachers can deliver high-quality instruction, to strengthen teachers' own understanding of literacy and numeracy concepts, and to improve academic leadership and school improvement planning by headteachers (USAID, 2014). It typically works through a two-year modular programme of workshops and school visits, after which schools continue to receive school visits from government officers to maintain and continue quality gains.

Under the fourth output stream, ESSPIN aims to improve inclusion practices and to strengthen community engagement in school improvement and wider access. In particular, Output stream 4 seeks to ensure that community members, including women and girls, influence the way schools are run; that community and government organisations are better able to press for school improvement; and that schools and communities ensure that the needs of all children are met. These interventions to improve community participation through functioning SBMCs come within a challenging sociocultural context. Qualitative research prior to the introduction of ESSPIN interventions (ESSPIN, 2009) found that community members were often not aware of SBMCs. SBMCs also lacked clarity on their roles and responsibilities, and lacked the resources to contribute effectively to school management. It was particularly difficult for women and students to participate, as this was a cultural taboo in many areas.

The programme's theory of change assumes that the interventions will improve five pillars (or outputs) of school quality: headteacher effectiveness, teacher competence, adoption of inclusive practices to meet the needs of pupils, introduction of SDPs, and establishment of functional school-based management. These pillars collectively contribute to an improvement in overall school quality (outcome), and this in turn increases pupil learning outcomes (impact).

Initially the programme was piloted in a sample of schools and managed by the ESSPIN infrastructure. As the programme was scaled up, management and delivery of the support (both

Output streams 3 and 4) came under the state governments. The state infrastructure then provided the training and mentoring, using the ESSPIN model and under guidance from ESSPIN staff. We continue to call the beneficiary schools 'ESSPIN schools' to indicate that they received the ESSPIN delivery model.

1.2 ESSPIN in Kaduna State

ESSPIN has been working in partnership with the government of Kaduna State to realise sustainable school improvement since 2009. In Kaduna the pilot (Phase 1) of the ESSPIN intervention focused on building the skills of teachers, headteachers and SBMCs drawn from members of the local communities (ESSPIN 2013b). From Phase 2, the intervention continued with the key areas of headteachers and SBMCs, but had a more targeted focus on improving teachers' ability to deliver stronger numeracy and literacy teaching through set lesson plans.

The key school-level interventions in Kaduna's schools partnering with ESSPIN were:

- Training for headteachers on:
  - academic leadership;
  - school planning;
  - management of teachers; and
  - working with the community.
- Training for teachers on:
  - generic basic teaching skills;
  - basic literacy teaching (initial reading skills);
  - basic numeracy teaching (number concepts, addition and subtraction); and
  - use of teaching aids, classroom organisation, and praise.
- Phase 1 schools received two school grants in consecutive years at an average of N150,000 per year (the exact sum depending on school size) to be spent on activities agreed by the headteacher and SBMC as priorities for school improvement and included in the SDP based on a school self-evaluation.
- 24 State School Improvement Team members (all lecturers at the College of Education in Gidan Waya) received ongoing training and support over a three-year period to develop the capacity of the Advisory Service Unit, 59 School Support Officers based in local government, and 25 District Support Officers (a layer of administration which exists only in Kaduna) to enable them to lead the school improvement process at school level.
- 26 civil society organisation members and government officers from the Department of Social Mobilisation received training on how to activate, train and mentor SBMCs.
- Community members (17 people from each ESSPIN school) received training on establishing an SBMC, which covered:
  - school planning and management;
  - SBMC roles and responsibilities;
  - communication and leadership;
  - women and children's participation in SBMCs;
  - resource mobilisation and financial processes;
  - inclusive education and gender;
  - child protection and participation; and
  - change and relationship management.

Annex A presents some descriptive statistics on the schools selected for ESSPIN and those not selected, while Annex B sets out the interventions under Output stream 3 made in Kaduna from 2009/10 through to 2013/14, indicating the number of days of training received by each headteacher, each teacher trained under ESSPIN, and the number of visits to the school. Annex C sets out the interventions under ESSPIN's Output stream 4, indicating the number of days of training for SBMCs, training on participation by women and children, and mentoring visits.

1.3 Selection of ESSPIN beneficiary schools and expansion

The ESSPIN programme has been gradually rolled out to more government primary schools in Kaduna, such that by time of the 2014 Composite Survey 27% of schools had been benefiting from the full package for at least one year (Table 1 and Annex B). The scale of the roll-out of Output stream 3 consisted of the following stages:

- The Phase 1 (pilot) began with 165 government primary schools (4% of the state's total) in 2009/10 and these schools continued receiving Output stream 3 activities (leadership training, teacher training and school visits) in all years up to the survey (ESSPIN 2013a). The schools selected for the pilot were all schools in two or three districts selected by the Education Secretaries of seven Local Government Education Areas (LGEAs) (two from the southern and central Senatorial District and three from the north).

- A Phase 2 roll-out began in 2011/12; these 317 schools were selected from the remaining LGEAs that had not been part of the State Education Sector Project.[1]

- The SIPs were further rolled out to new schools from 2012/13 and then 2013/14.

Once a school was in the programme, the provision of leadership training and school visits was continuous every year (although not all beneficiary schools received teacher training in 2013/14). Therefore the more years a school had been in the programme, the higher the expected improvement in output, outcome and impact indicators.

Table 1. Proportion of schools receiving full package of ESSPIN Output stream 3 interventions

%	2009/10	2010/11	2011/12	2012/13	2013/14	Any year
Kaduna	4	4	14	21	13	27

Source: Authors' calculations based on 2012/13 annual school census and intervention information provided by ESSPIN.
Note: Proportions are calculated relative to the total number of schools in the 2012/13 annual school census, and so these are not perfectly accurate for other years because the total number of schools changes slightly from year to year. Where census numbers are lower than ESSPIN's intervention tables, the information from ESSPIN is used on the assumption that there is some missing data in the school census

The expansion of the programme to more schools in Phase 2 required a changed model for delivering training, with state governments taking on the management, and the training located closer to schools. The change in model makes delivery cheaper per school and more sustainable for the states to run themselves, as well as enabling states to take control, all of which were necessary to allow scale-up. Programme staff argue that locating training closer to the schools has longer term benefits. However, we might see that in the shorter term quality standards from the pilot programme are not fully upheld as the new, much larger numbers of trainers, who typically have lower qualifications than the first wave, develop competencies. The survey results also show that there is high variation in the indicators among schools in Kaduna. A heterogeneous group of

[1] The State Education Sector Project was a World Bank-supported project to improve the quality of basic education in targeted Local Government Areas through a school development scheme. The project closed in 2011.

schools means there will be differing needs and demands, and this is likely to make designing and running a SIP more challenging.

A summary of the characteristics of Kaduna's schools according to the level of ESSPIN intervention is given in Annex A. Schools with more years of ESSPIN intervention were generally more likely to be urban and have higher enrolment and more teachers, and a lower pupil–teacher ratio (PTR).

Enrolment in Kaduna's government primary schools, as captured in the census, increased by more than 18%, rising from 970,000 to 1,150,000 between 2009 and 2013 (Table 2). Enrolment in the schools captured in both censuses increased by 16%. The increase in enrolment was accompanied by increases in PTRs across all schools (Annex A). Average PTRs of around 30–45 hide substantial variation, with some schools and some grades experiencing very high numbers of pupils per classroom and per teacher.

Table 2. Number of schools and enrolment in the 2009 and 2013 school censuses

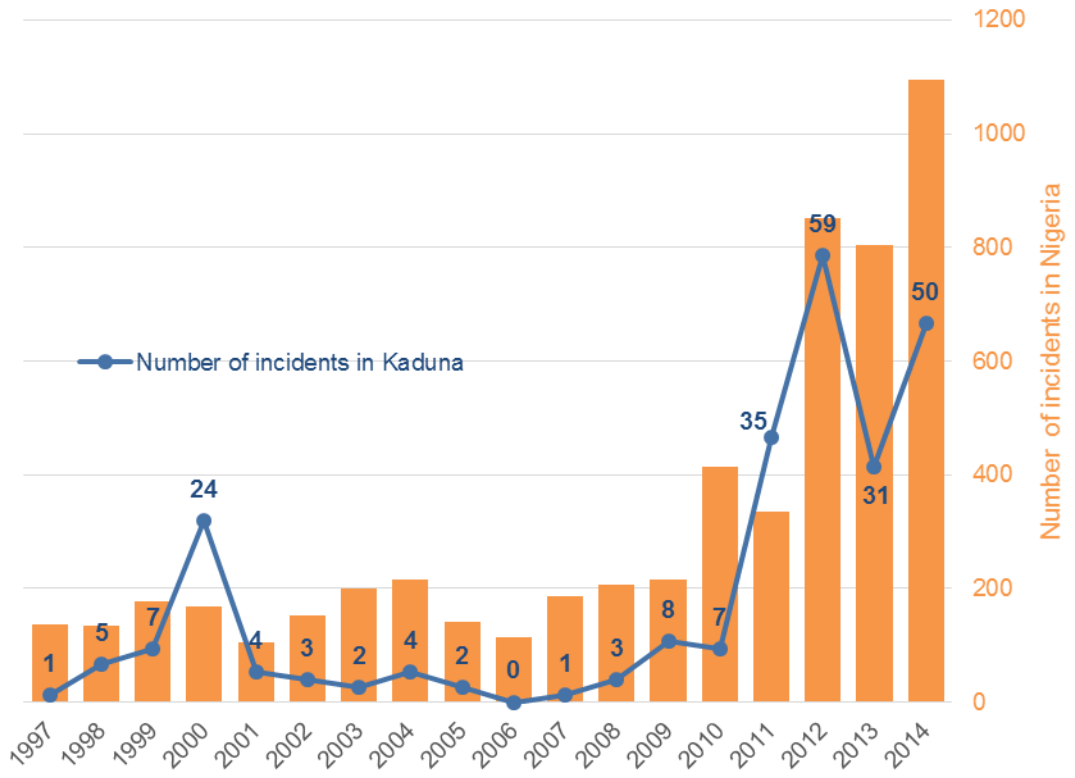
	2009		2013		Enrolment change (%)	Enrolment change (schools found in both censuses only, %)
State	Schools	Enrolment	Schools	Enrolment		
Kaduna	3947	972,985	4223	1,151,876	18.4	16.2

Note: Enrolment is for primary grades 1–6.

1.4 Conflict in Kaduna

This report is written in the context of growing insecurity in Nigeria, particularly in, but not limited to, three states of the north-east in which a state of emergency has been declared (Borno, Yobe and Adamawa). Across Nigeria the number of recorded incidents of political violence and conflict has increased eightfold since 1997. In Kaduna the incidents of violence have also increased in this time period, most notably since 2010 (Figure 1). In 2014, there were 50 recorded violent events in Kaduna, causing 538 fatalities (Table 3).

Figure 1. Incidents of political violence in Nigeria and Kaduna



Source: Armed Conflict Location & Event Data Project (ACLED), Version 5 (1997–2014). Note all events from ACLED are included except for those categorised as protests which did not involve a fatality.

Table 3. Kaduna: Political violence: Incidents and fatalities, 2010–2014

Variable	2010	2011	2012	2013	2014
Events	7	35	59	31	50
Fatalities	20	851	269	140	538

Source: ACLED, Version 5 (1997–2014). Note all events from ACLED are included except for those categorised as protests which did not involve a fatality.

ESSPIN worked with Kaduna, Kano and Jigawa in September 2014 to conduct small-scale, in-depth research into the impact that conflict and insecurity may be having on education (six schools per state), and the way in which education may impact on or fuel conflict in Nigeria (ESSPIN 2014). Although no state of emergency has been declared in these states, they have all been affected in different ways by conflict and violence, including both communal violence and insurgency.

Evidence from the 18 school communities suggests that there is a growing impact on schools and communities in ESSPIN-supported LGEAs in Kaduna, which should be taken into account in the reading of this report.

As a specific example of local conflict disrupting education, during the data collection for CS1 a period of civil unrest in northern Nigeria was felt very strongly in Kaduna (ESSPIN 2013b). A curfew period made travel unsafe and postponed the data collection for two weeks, after which the survey method had to be adjusted (with a team reallocation, and shortening the time spent in schools) in order to make it safe for enumerators, and fieldwork spilled into the last week of term, which is not usually spent in lessons. We are not able to quantify the impact on pupils, teachers, communities and the data collection, but it must have been considerable.

2 Methodology and analysis

2.1 Evaluation strategy

2.1.1 Classifying the amount of ESSPIN intervention

For the purposes of evaluation, ESSPIN was originally intended to be rolled out in a simple phased pattern across the six states, with schools falling into one of three groups: no intervention (control), Phase 1 (roll-out prior to the 2012/13 school year), and Phase 2 (roll-out in 2012/13 or 2013/14). In practice, Kaduna State decided to extend the programme over four phases (the original 2009/10 pilot, a 2011/12 group, 2012/13 group, and lastly a 2013/14 group), based on the government's capacity and willingness. The majority of schools in Kaduna, as of 2013/14, have not received school-level ESSPIN intervention, and so remain part of the control group.

For the purposes of evaluation, we have grouped schools according to the number of years of the 'full package' of Output stream 3 support they have received (see Annex B for full details).[2] We then have two main categories of schools: ESSPIN schools and non-ESSPIN schools. The non-ESSPIN, or 'control schools', include those which received no ESSPIN Output stream 3 intervention at all as well as the 200 schools that had support only in 2013/14, considered too recent for the intervention to have taken effect. Table 4 in the following section shows the number of schools in each of these categories. Note that 'ESSPIN schools' is used as a shorthand for government schools which have benefited from the ESSPIN model of school improvement, whether delivered by the state government or the ESSPIN programme staff.

When we are looking at one point in time (cross-sectional analysis), the ESSPIN schools, or 'treatment schools', are expected to be better because of ESSPIN. When we are looking at change over time, the ESSPIN schools are expected to have improved faster because of ESSPIN.

For individual outcome indicators, we alter the classification scheme slightly according to the purpose. For example, for examining teacher competence, we consider three different groups: teachers that have not been exposed to ESSPIN; teachers who are in schools that have received ESSPIN intervention but who have not themselves been trained through ESSPIN; and teachers who have been trained through ESSPIN. We also use continuous versions of the intervention measures – for example, the number of years that a pupil has been exposed to expected improved school quality as a result of ESSPIN intervention. While categorical measures are easier to use for tables of descriptive statistics, a continuous measure makes sense in regression analysis, makes most use of the information, and helps us to avoid the risk that results might be altered by a slight change in the choice of categories.

2.1.2 Modes of analysis

The purpose of CS2 is both to provide insight into the changes over time in the six states where ESSPIN works, and to evaluate whether the ESSPIN model is having an effect in the specific schools where its school improvement and community inclusion interventions have operated. We are interested in a wide range of output indicators: teacher competence, headteacher effectiveness, school development planning, school inclusiveness, and the functionality and inclusiveness of SBMCs. Some of these same indicators are also combined to give an overall indicator of school quality. Finally, ESSPIN's impact is measured in terms of improved pupil

[2] A companion report, *Composite Survey 2: Gender and Inclusion Report* (De and Cameron, 2015), focuses on ESSPIN's Output stream 4 interventions, which run in parallel with Output stream 3 and aims to improve inclusion and community participation in schools.

learning outcomes, which we ascertain through test scores in numeracy and English literacy in grades 2 and 4. For each of these indicators, we present in the following chapter three main types of analysis:

1. Change over time between CS1 and CS2, for Kaduna as a whole. These changes likely reflect changes that are beyond the control of ESSPIN. Although the recent expansion of ESSPIN interventions has meant that the programme now has direct links with a larger number of schools in Kaduna, much of this roll-out happened in 2013/14 and so is unlikely to have started having a major impact by the time of our survey, near the end of the 2013/14 school year.

2. Differences between ESSPIN and non-ESSPIN schools within the CS2 results. In the ESSPIN schools we hypothesise that our output, outcome and impact measures will all be higher than in the control group. If this is the case, it provides good initial evidence that ESSPIN is effective, although it does not rule out the possibility that ESSPIN schools' better results could come from differences in school background characteristics pre-dating the ESSPIN intervention.

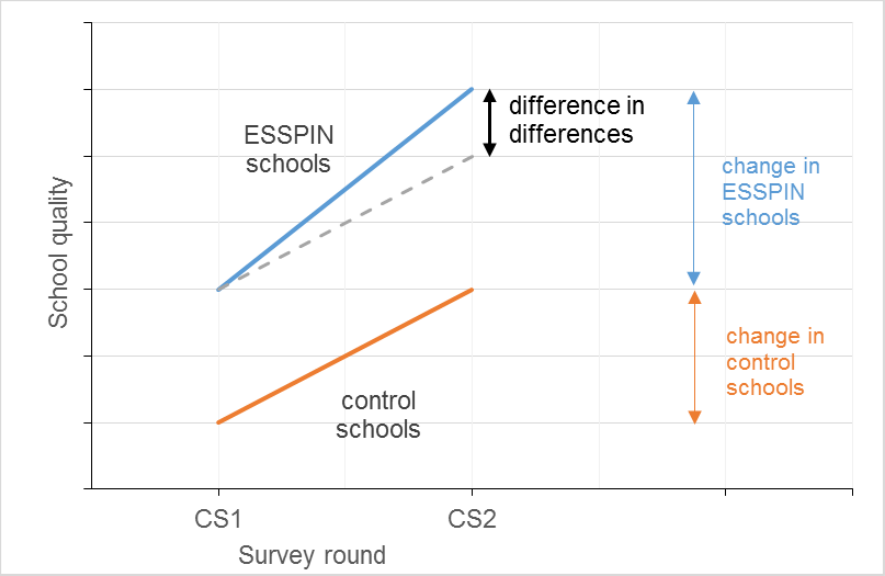
3. Difference in differences between ESSPIN and non-ESSPIN schools and over CS1 and CS2. We hypothesise that ESSPIN schools are expected to have improved faster than non-ESSPIN schools between 2012 and 2014. See Box 2 below.

In each case we use statistical significance tests (t-tests or z-tests) to give an indication of whether a difference in results (over time or between intervention groups) is significant. This should not be taken as rigorous hypothesis testing (given the very large number of indicators tested), but it provides a guide to whether a difference between the weighted average results in two groups is large enough relative to the variance of the results, to be able to provide us with a useful indication of likely differences in the population of schools in the six states. For analysing difference in differences we also use regression analysis; these are reported in Annex D.

Box 2. Difference in differences

The Composite Survey may reveal that ESSPIN schools are of higher quality, or have better learning outcomes, than other schools. But how do we know whether this can be attributed to ESSPIN and is not just because ESSPIN schools were better in the first place? One way is to focus on change over time using 'difference in differences' methods. The underlying idea is that schools which have had ESSPIN interventions between CS1 and CS2 – that is, between 2012 and 2014 – ought to have improved faster during that period than schools which did not have ESSPIN interventions.

We can measure this degree of improvement by comparing averages of the indicator of interest – school quality, say – during CS1 and CS2, in control schools and ESSPIN schools. Is the change over time greater in the ESSPIN schools than in non-ESSPIN schools? If so – and if statistical tests confirm that this result is unlikely to have occurred by chance – then this is considered good evidence that ESSPIN itself had an effect and was not just lucky in choosing schools that were good in the first place (selection bias).



Does a significant difference in differences (or treatment effect) prove that the faster improvement in some schools can be attributed to ESSPIN? Not absolutely. It is still possible that there are other factors at play causing a faster improvement in some schools than others. For this reason, in the overall CS2 report we use other statistical techniques to examine whether ESSPIN schools had different characteristics to start with, and to control for any such differences.

2.2 Sampling, coverage and weights

In Kaduna the sample allocation for CS1 was 105 schools, giving 35 each in control, Phase 1 and Phase 2 categories of schools (as they were categorised at the time). There was a large amount of variation pupil and teacher indicators between schools in the state, which reduced the precision with which we could estimate these indicators. In order to improve the precision of our estimates we increased the sample of schools in Kaduna to 140 for CS2, following the recommendations of an analysis of the issue in CS1 (Megill, 2014). The number of schools sampled in each of the categories (as defined in CS2, so taking account of the full period of intervention) is shown in Table 4.

Table 4. Sample in CS1 and CS2 and population of schools, by intervention group

	Category for sampling purposes	CS1 sample	CS2 sample	Population	Categories for analysis	
					Expected to be better at CS2	Expected to have improved between CS1 and CS2
Kaduna	none/minimum	28	61	3413	no	no
	medium	42	42	671	yes	yes
	maximum	35	37	165	yes	yes
	Total	105	140	4249		

Note: The sample size shown is the actual sample for which data was collected.

In each school the headteacher was interviewed, as was the SBMC chairperson.

Teachers within each school were sampled from the population present in the school on the day of the survey visit and who taught grade 1–6 in the present term, using the school's teacher attendance register. The sample was reduced from 10 teachers in CS1 to six teachers in CS2 in order to improve the accuracy of the indicators.

Pupils were sampled from the pupil registers for grade 2 and 4 classes – four each for numeracy and literacy by grade.

Within the schools, it was not always possible to administer all of the intended instruments. This could happen because the school was very small, and lacked a sufficient number of pupils and eligible teachers. It also sometimes happened that teachers and pupils were not present at 8am, when sampling was conducted; and occasionally pupils and teachers would leave the school after being sampled (for example, due to illness). In total 92% of the intended sample of pupils was included, and 70% of teachers. The actual numbers of schools, teachers and students sampled is given in Table 5.

Table 5. Kaduna: Sample coverage in CS2

	Schools		Teachers			Pupil tests			
	Intended sample	Actual	Interview	Less. Obs.	Tests	L2	L4	N2	N4
Kaduna	140	140	638	625	586	539	521	538	517

Note: *L2* refers to the grade 2 literacy test, *L4* to the grade 4 literacy test, *N2* to the grade 2 numeracy test, and *N4* to the grade 4 numeracy test.

Simple averages of the results from the Composite Survey data would not be representative of what is happening across the state, because (as Table 4 above shows) in terms of the proportion of schools in each of the roll-out phases, the profile of schools in the survey is not identical to the profile of schools in the state as a whole. We overcome this by applying sample weights which give greater weight to the results in schools that are relatively under-represented in the survey. Sample weights were calculated for the CS1 and CS2 schools, teachers and pupils.

2.3 Fieldwork and instruments

Fieldwork for CS2, including the pupil tests, was conducted during May–July 2014. The following data collection was carried out:

(i) Structured interviews were conducted with teachers, headteachers, and SBMC chairpersons;

(ii) A lesson observation was conducted for each teacher sampled;

(iii) Teacher tests were conducted at the end of the survey, in a number of testing centres in each state; and

(iv) Pupils in primary grades 2 and 4 were given tests in either literacy or numeracy.

Two indicators of aggregate learning outcomes are used in this report. The first is the total mark achieved by the pupil in each test paper, expressed as a percentage score. The second is the proportion of tested pupils who successfully answer a subset of questions which aim to measure a specific field of learning, as described in ESSPIN's logframe. Although the latter may be important for assessing ESSPIN's success in improving specific types of learning (e.g. the ability to read with comprehension), their reliance on data from a small number of questions (2–3) is statistically problematic. They are less reliable and less sensitive indicators than the total mark, which uses all the data available. For completeness, both types of indicator are used in this report.

3 Findings

Box 3. How to interpret the analysis, and expected results

For each indicator, three types of analysis are presented:

- Comparison of averages between CS1 (2012) and CS2 (2014). Here the results are representative of all schools (or teachers, or pupils) in the state, as found in CS1 and then in CS2. This depends on both general trends at the state level and any improvements in ESSPIN schools depending on the scale of ESSPIN roll-out. The hope is that ESSPIN state-level interventions combined with the SIP will lead to an improvement in state-wide averages.
- Comparison of groups in CS2, according to whether they have benefited from school-level ESSPIN intervention or not. Here we expect the results to be better for schools which have benefited from ESSPIN. If a school entered the programme in 2013/14, we count it as non-ESSPIN, as we would not expect the support to have impacted on the indicators yet.
- Comparison of schools which benefited from support in 2011/12 and 2012/13, with those which did not, to see whether they improved more or faster between 2012 and 2014. Here we expect the supported schools to improve relative to other schools over the two years.

Many of the statistics estimated for Kaduna are characterised by high variance within periods or within groups of schools. This variance means that it is more difficult to detect statistically a significant effect in the data, such as a difference between ESSPIN and control groups or a change over time. Due to this, some of the indicators show differences that are substantial in magnitude yet not statistically significant, making it harder to draw strong conclusions. Where a pattern of results is consistent across different indicators, and large in magnitude, we consider this inconclusive but indicative evidence of a real effect in the population, even if the difference does not reach statistical significance.

3.1 Teacher competence

3.1.1 Main analysis

The ESSPIN logframe sets four criteria for judging competence of teachers (Box 4). A teacher who teaches English or maths is defined as competent if he or she meets at least three of these, while teachers of other subjects are exempted from one of the four criteria (knowledge of the English or maths curriculum) and defined as competent if they meet two of the remaining three criteria.

For CS2, a fifth criterion was added, based on teacher test results. Teachers are defined as competent if they are competent according to the original criteria, and can also score at least 50% in primary school-level literacy and numeracy tests.

Box 4. Logframe standard for teacher competence

A teacher must meet three out of four of the following criteria to meet the competence standard if he/she teaches English and/or maths. Teachers of other subjects must meet two out of three criteria (excluding 1 below):

1) Knowledge of English or mathematics curriculum (based on interview);
2) Use of at least one teaching aid during lesson observation;
3) Greater use of praise than reprimands during lesson observation; and
4) In terms of class organisation: assigning individual or group tasks at least twice during lesson observation (or for two contiguous five-minute blocks).

For CS2, a new stricter indicator of teacher competence has been introduced. This excludes reading from or writing on, or having pupils copy from, the blackboard as a use of a teaching aid, and adds a fifth criterion:

5) Literacy and numeracy: scores at least 50% in both an English literacy and a numeracy test.

Table 6 compares the results for Kaduna's teachers in CS1 and CS2. It shows the average for all teachers in Kaduna, balancing the results of teachers who did and did not benefit from ESSPIN in order to make it representative of the average. (The fifth criterion is not available here as teacher tests were not conducted as part of CS1.) Generally, the level of teacher competence appears to have fallen in Kaduna according to the two surveys. Teachers' knowledge of the English or maths curriculum has dropped in Kaduna from 63% to 31%. The proportion of teachers using teaching aids and praise in the classroom increased, although this was not statistically significant. Overall in Kaduna, the proportion of teachers meeting the (CS1) competence standard is now 62%, which is a significant fall from 76% in CS1.

We also calculate a continuous 'competence score', based on the number of criteria met by each teacher. A teacher who meets all of the three or four criteria would score 100%, while a teacher who meets none of them would score 0%. The competence score is lower in CS2 than in CS1. This drop in the results for teacher competence may be due to some measurement error in the first round of the survey. It was noted in the CS1 report that performance was surprisingly strong given the weak scores for teachers found in an earlier (2010) Teacher Development Needs Assessment (ESSPIN, 2013b). It is possible that CS1 over-estimated the indicators of teacher competence. This could have happened, for example, if data collectors administered the question about English and mathematics curriculum benchmarks wrongly.

Table 6. Kaduna: Teacher competence in CS1 and CS2

	CS1	CS2	
(1) Knowledge of Eng/maths curriculum	62.9	30.5	-
(2) Use of 1+ teaching aid	94.9	95.8	
(3) Praise more than reprimand	76.4	79.4	
(4) Assigns 2+ ind./group task	60.2	50.8	
Competence score (CS1 version)	75.9	66	-
Teacher competence standard (CS1)	76.2	62.3	-

Note: + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

Focusing on the findings in CS2, the direction of results suggests that teachers perform better in ESSPIN schools than in non-ESSPIN schools (Table 7). We distinguish three groups of teachers: (1) those who are in schools that received no ESSPIN intervention; (2) those who are in schools

that received ESSPIN intervention but who did not individually receive ESSPIN teacher training; and (3) those who are in ESSPIN schools and individually received ESSPIN teacher training. Generally the proportion of teachers meeting the indicators was higher for teachers who had been trained than those who had not, and two of these reach statistical significance: the use of teaching aids (other than the blackboard), and the use of praise more than reprimands in the classroom. The results for teachers in ESSPIN schools who had not been trained themselves are generally better than for teachers in non-ESSPIN schools, particularly in terms of curriculum knowledge, using teaching aids other than the blackboard, using praise and assigning varied tasks. Although this positive difference was not found to be significant, it may reflect some positive spillover effects from the trained teachers, who are encouraged to share their learning with colleagues. The overall competence score was higher (though not statistically so) for teachers who had been trained under ESSPIN and those not trained but in ESSPIN schools, compared with teachers in non-ESSPIN schools.

Table 7. Kaduna: Teacher competence in CS2, by intervention group

	(1) Non-ESSPIN	(2) ESSPIN school		(3) ESSPIN-trained	
Knowledge of Eng/maths curriculum	26.9	39.7		37.4	
Use of 1+ teaching aid	95	97.6		98	
Use of 1+ teaching aid excl read/write/copy from blackboard	75.8	90.5		93.7	+
Praise more than reprimand	77	81.4		94.2	+
Assigns 2+ ind./group task	48.6	54.5		59.5	
Literacy score (%)	45.8	44.2		50.5	
Numeracy score (%)	59.4	59.2		60.7	
Passes literacy and numeracy test	36.6	31.5		48.9	
Competence score (CS1 version)	63.8	70.4		73.3	
Teacher competence standard (CS1)	58.1	72.3		70.8	
Competence score (CS2 version)	59.2	64.4		68.4	
Teacher competence standard (CS2)	20.7	19.6		35.4	

Note: The CS2 version of the competence score adds the teacher's performance in the literacy and numeracy tests to the number of other criteria met by the teacher; for example, a teacher who met all four original criteria and also scored 100% in the literacy and numeracy tests would receive a competency score of 100%; + indicates a significant difference from the results in non-ESSPIN schools.

Did teachers benefiting from ESSPIN interventions improve faster than those who did not between 2012 and 2014? The comparison of means (Table 8) finds there was no significant difference between the change in competence for non-ESSPIN and ESSPIN-trained teachers. The general trend suggests a fall in teachers' average competence scores, with the group of teachers not in ESSPIN schools (column 1) seeing the largest fall in scores. The groups of teachers in ESSPIN schools (column 2) and those individually trained by ESSPIN (column 3) were less adversely affected, but the difference from the non-ESSPIN school teachers was not significant. An alternative method of analysing difference in differences, using regression (see Annex D), finds similar results.

Table 8. Kaduna: Teacher competence difference in differences (comparison of means)

Teacher competence scores (CS1 version)	(1) Non-ESSPIN	(2) ESSPIN school	(3) ESSPIN-trained
CS1	76.7	69.9	81.9
CS2	63.8	70.4	73.3
Difference	-12.9	0.5	-8.6

Note: * indicates a significantly different difference than that in non-ESSPIN schools (p < .05).

Generally, teachers who have received training through ESSPIN were more likely to meet some competence criteria than those without training in 2014. However, there is no strong evidence to suggest that receiving individual training through ESSPIN in recent years improved the competence of teachers in Kaduna relative to other teachers. The findings suggest that there was no improvement in the average level of teacher competence in Kaduna between 2012 and 2014.

3.1.2 Findings from the teacher content knowledge tests

The teacher tests included items pitched at primary school grades 1 to 5 and focusing on different areas: foundational skills for teaching literacy; writing; reading; grammar; number concepts; calculation; and other numeracy skills. In Kaduna, teachers were twice as strong in reading as they were in writing, and grammar was also a stronger area (Figure 2). In mathematics, teachers performed better on number concepts than calculation and other numeracy items. As would be expected, teachers' ability to answer the questions falls as the grade level of the questions increases (Figure 3). On the whole the mathematics items were easier for teachers in Kaduna than the English items.

Figure 2. Kaduna: Teacher test scores across domains of learning

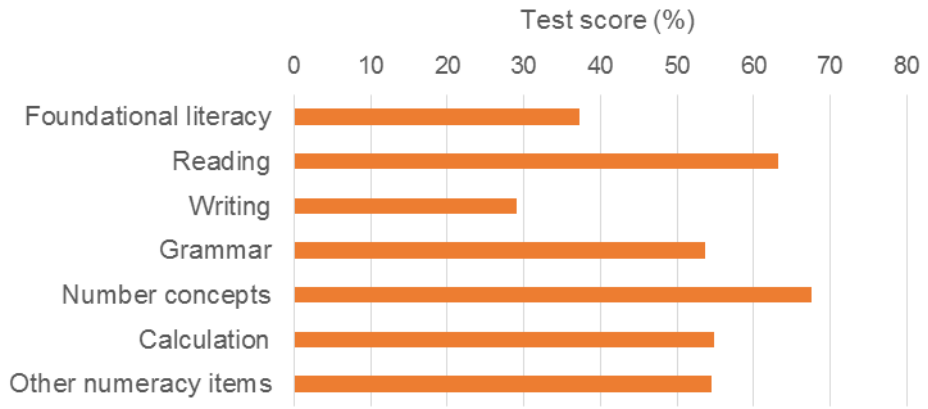
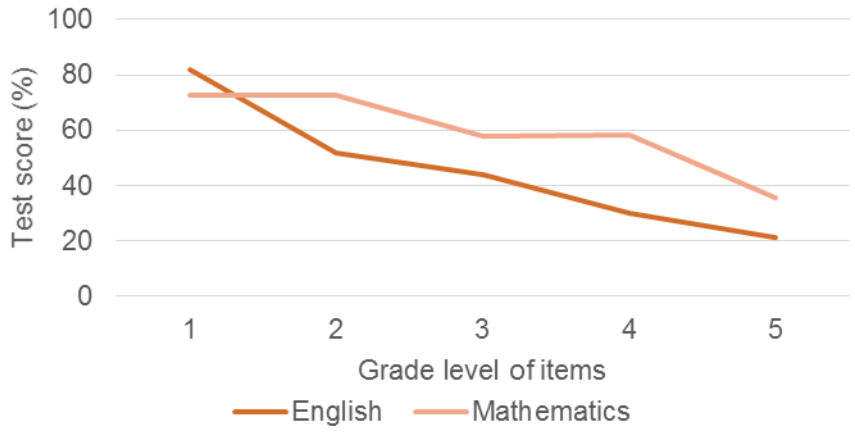


Figure 3. Kaduna: Teacher test scores by grade



3.2 Headteacher effectiveness

The ESSPIN logframe defines headteacher effectiveness in terms of seven criteria (Box 5). These reflect both activities by the headteacher and behaviour across the teachers and pupils such as agreement on what time the school should open (criterion 4), presence in class at the beginning of the school day (criterion 5), and appropriate break and lesson durations (criteria 6 and 7).

Box 5. Logframe standard for headteacher effectiveness

A headteacher must ensure that five out of seven of the following criteria are met in order to meet the headteacher effectiveness standard:

1) Carry out two or more lesson observations in the past two weeks;
2) Hold four or more professional development meetings since the start of the 2011/12 or 2013/14 school year (NB: the survey took place more than nine months into the school year);
3) School has a teacher attendance book and the headteacher recalls at least two actions taken to promote teacher attendance;
4) Clear school opening time: more than 50% of pupils sampled agree on the school opening time and more than 50% of teachers sampled agree on the school opening time;
5) More than 50% of classes are in their classroom with their teacher within 30 minutes of school opening time;
6) Length of morning break is 35 minutes or less, except in Enugu when it must be 15 minutes or less; and
7) More than 50% of lessons observed finished within five minutes of a standard 35-minute lesson duration (i.e. the lesson was between 30 and 40 minutes long).

The proportion of all headteachers who met the criteria for effectiveness rose from 8.5% to 13.2% between CS1 and CS2 in Kaduna (Table 9), but the change was not statistically significant. The proportion of heads meeting the individual criteria remained much the same, except for a significant reduction in the proportion who took action to improve teacher attendance; there was also an improvement in schools with a clear opening time and appropriate lesson lengths (although the latter changes were not statistically significant).

Table 9. Kaduna: Headteacher effectiveness in CS1 and CS2

	CS1	CS2	
(1) Lesson observations	4.6	2	
(2) Professional development meetings	10.1	4.3	
(3) Action on teacher attendance	88.2	58.7	-
(4) Clear opening time	49.8	61.1	
(5) In class on time	77.2	72.3	
(6) Appropriate morning break	85.4	83.7	
(7) Appropriate lesson length	26.6	46.1	
Number of criteria fulfilled (/7)	3.5	3.4	
Effective headteacher (5/7 criteria met)	8.5	13.2	

Note: + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

In the 2014 data, headteachers who have received training are significantly more likely to be conducting lesson observations and professional development meetings than those who have not (Table 10). The proportion of schools who met the effective headteacher indicator (five out of seven sub-criteria), and the average number of criteria fulfilled, were lower in ESSPIN schools than non-ESSPIN, but the differences were not statistically significant. Overall, we therefore have limited evidence to suggest that head teachers are more effective in ESSPIN schools.

Table 10. Kaduna: Headteacher effectiveness in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) Lesson observations	0.2	9.7	+
(2) Professional development meetings	0.6	19.9	+
(3) Action on teacher attendance	60.1	52.7	
(4) Clear opening time	65.3	43	
(5) In class on time	71.9	74.1	
(6) Appropriate morning break	85.2	76.9	
(7) Appropriate lesson length	49.2	33.3	
Number of criteria fulfilled (/7)	3.4	3.1	
Effective headteacher (5/7 criteria met)	13.3	12.7	

Note: + indicates a significant difference from the results in non-ESSPIN schools.

As in the previous section on teacher competence, we also examine change over time in headteacher effectiveness to see whether headteachers in schools that received more ESSPIN intervention between 2012 and 2014 improved faster (or in this case declined more slowly) than comparators. The results suggest there was no significant difference in the change in number of criteria met by headteachers in schools benefiting from ESSPIN and schools which did not expect an improvement (Table 11). We also adjusted for the year that the headteacher was appointed to his or her current school to take account of any transfers, and still found no evidence of a difference between the two groups. A regression analysis found similar results and is reported in Annex D.

Table 11. Kaduna: Headteacher effectiveness difference in differences (comparison of means)

Number of criteria met (/7)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	3.6	3.2
CS2	3.4	3.1
Difference	-0.2	-0.1

Note: * indicates a significantly different difference than that in non-ESSPIN schools (p < .05).

In summary, there is no evidence that headteacher effectiveness has changed significantly overall in Kaduna, although there is some evidence of an improvement in specific criteria. The headteachers who are expected to have improved due to leadership training under ESSPIN have better performance than other headteachers on two out of seven criteria. The results do not support the expectation that headteachers benefiting from leadership training would have improved faster than other headteachers between 2012 and 2014.

3.3 School development planning

The definition of effective school development planning depends on five criteria (Box 6). In Kaduna there was no significant change in average school performance on the five criteria between CS1 and CS2, although some criteria increased and others decreased (Table 12). Overall, it is hard to say whether school development planning has changed in Kaduna in the past two years.

Box 6. Logframe standard for effective school development planning

The school must meet criterion 1 and criterion 2 listed below and at least two out of three of the remaining criteria in order to meet the effective school development planning standard:

1) Written evidence of school self-evaluation process for current school year;
2) SDP for current school year available;
3) SDP contains three or more activities which aim to strengthen teaching and learning;
4) Physical evidence of four or more activities from SDP having been carried out; and
5) Cashbook is up-to-date (balanced in the last 60 days).

Table 12. Kaduna: SDP effectiveness in CS1 and CS2

	CS1	CS2	
(1) Written evidence of school self-evaluation process	26.3	10.8	
(2) SDP available	19.5	11.3	
(3) SDP contains 3+ activities to strengthen teaching and learning	3.4	2.4	
(4) Evidence that 4+ activities from SDP carried out	1.4	2.9	
(5) Cashbook up-to-date	14.2	25.1	
Number of SDP criteria fulfilled (/5)	0.6	0.5	
School meets effective school development planning standard	1.4	2.1	

Note: + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; p < .05),

Focusing on CS2, ESSPIN schools are significantly better at carrying out school self-evaluation and having a current SDP (Table 13). The results also show a positive difference for activities to improve teaching and learning, and in the number of activities that were actually carried out, but

the differences are not statistically significant. The average number of criteria fulfilled is higher for ESSPIN schools at 1.3 out of 5, compared with 0.4 on average for other schools, and this was a significant difference. While 11% of ESSPIN schools met the overall standard, none of the non-ESSPIN schools did.

Table 13. Kaduna: SDP effectiveness in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) Written evidence of school self-evaluation process	4.9	35.9	+
(2) SDP available	3.9	43.2	+
(3) SDP contains 3+ activities to strengthen teaching and learning	0	12.6	
(4) Evidence that 4+ activities from SDP carried out	0	15.5	
(5) Cashbook up-to-date	26.4	19.3	
Number of SDP criteria fulfilled (/5)	0.4	1.3	+
School meets effective school development planning standard	0	11	

Note: + indicates a significant difference from the results in non-ESSPIN schools.

Difference in differences analysis to assess whether ESSPIN schools improved faster than non-ESSPIN schools in terms of school development planning found no significant difference in the change between ESSPIN and non-ESSPIN schools (Table 14). A method using regression analysis gives the same findings (Annex D). This means we cannot say that the ESSPIN interventions made schools more likely to improve in school development planning in this time period.

Table 14. Kaduna: SDP effectiveness difference in differences (comparison of means)

Number criteria met (/5)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	0.5	0.9
CS2	0.4	1.3
Difference	-0.2	0.4

Note: * indicates a significantly different difference than that in non-ESSPIN schools ($p < .05$).

In summary, school development planning appears to have remained at a low level in Kaduna with 2% of all schools meeting the standard in 2014, and is better for the sub-group of schools which benefited from ESSPIN activities (11% met the standard). There is no evidence to suggest that school development planning in the schools which had more ESSPIN intervention improved faster than schools without the intervention between CS1 and CS2.

3.4 School inclusiveness: meeting the needs of all pupils

The school inclusiveness standard depends on meeting three out of four criteria (Box 7), and schools are defined as partially meeting the standard if two criteria are met. In Kaduna generally there was an improvement in the proportion of schools meeting two of the criteria between 2012 and 2014, and a decline in two others, although no changes were statistically significant (Table 15). The proportion of schools meeting the standard has fallen, but not significantly. The reductions led to a statistically significant fall in the average inclusiveness score – from 72% to 62%.

Box 7. Standard for school inclusiveness (meeting needs of all pupils)

The school must meet at least three of the four criteria listed below in order to meet the school inclusiveness standard. The standard is partially met if two criteria are met:

1) Headteacher states three or more actions that he/she has taken to improve pupil attendance;
2) SDP contains two or more activities which aim to improve access;
3) More than 50% of teachers observed provided evidence of using two or more assessment methods (marked class test, marked pupil workbook, or graded examination paper); and
4) More than 50% of teachers observed met the spatial inclusion criterion (defined as engaging with at least one pupil from four different areas of the classroom during a lesson) and more than 50% of teachers observed met the gender inclusion criterion (defined as engaging with boys and girls proportionally to their presence in the classroom within a 10% margin; for example, if the class contains 50% girls then teachers who engage with girls between 60% and 40% of total engagements meet the criterion).

Table 15. Kaduna: School inclusiveness in CS1 and CS2

	CS1	CS2	
(1) 3+ actions to improve attendance	48.3	50.6	
(2) 2+ activities in SDP to improve access for disadvantaged children	3.6	5.3	
(3) >50% of teachers use 2+ assessment methods	71.3	55.1	
(4) >50% of teachers spatially inclusive and >50% are gender inclusive	38.4	28.4	
Number of inclusiveness criteria fulfilled (/4)	1.7	1.3	
Inclusiveness score	71.6	61.7	-
School partially met inclusiveness standard (2–4 criteria out of 4)	54.9	47.3	
School fully met inclusiveness standard (3–4 criteria out of 4)	23.1	11.2	

Note: + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

In 2014 ESSPIN schools performed better than non-ESSPIN schools on three of the four indicators, although the differences were not found to be statistically significant (Table 16). The number of ESSPIN schools meeting the overall standard was 21%, higher than in other schools (9%), but this was not significant. Schools in the ESSPIN programme did take part in more activities to improve access for disadvantaged children than non-ESSPIN schools.

Table 16. Kaduna: School inclusiveness in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
Inclusiveness criteria			
(1) 3+ actions to improve attendance	54.6	33.7	
(2) 2+ activities in SDP to improve access for disadvantaged children	2.5	17.3	
(3) >50% of teachers use 2+ assessment methods	50.1	76.9	
(4) >50% of teachers spatially inclusive and >50% are gender inclusive	27.4	32.8	
Overall inclusiveness standard			
Number of inclusiveness criteria fulfilled (/4)	1.3	1.6	
Inclusiveness score	61.2	63.9	
School partially met inclusiveness standard (2–4 criteria out of 4)	46.3	51.6	
School fully met inclusiveness standard (3–4 criteria out of 4)	8.8	21.2	
Detailed			
Number of actions to improve attendance	2.5	1.9	
Number of activities on access for disadvantaged children	0.1	0.6	+
Average number of assessment methods used	1	1.3	
Average number of zones participating in lessons	3.2	3.9	
Average gender equity score (0=completely unequal, 100=perfectly equal)	86.3	83.8	

Note: The gender equity score for a teacher is $100 - 100 \times abs(\frac{g}{g+b} - \frac{G}{G+B})$ where *g* is the number of girls who participate, *b* is the number of boys who participate, *G* is the number of girls present in the class, and *B* is the number of boys present in the class. It is expressed as a percentage score. For a lesson where the proportion of girls and boys participating is exactly equal to the proportion of girls and boys sitting in the lesson, the gender equity score will be 100; for a lesson where no boys participate, or no girls participate, the score will be zero. + indicates a significant difference from the results in non-ESSPIN schools.

Difference in difference analysis of schools in Kaduna depending on the level of ESSPIN intervention has been carried out by comparing the change in average inclusiveness score (Table 17). The results suggest that both groups of schools became less inclusive between 2012 and 2014, and by around the same amount. There is no evidence to support the suggestion that ESSPIN schools improved (or declined) at a different rate to schools not expected to benefit from the programme. A regression analysis finds the same results in Annex D.

Table 17. Kaduna: School inclusiveness difference in differences (comparison of means)

Inclusiveness score	(i) Non-ESSPIN	(ii) ESSPIN
CS1	71	73.1
CS2	61.2	63.9
difference	-9.8	-9.2

Note: * indicates a significantly different difference than that in non-ESSPIN schools ($p < .05$).

3.5 SBMC functionality and inclusiveness

ESSPIN conducted qualitative research into SBMCs and community engagement in education in five ESSPIN states in 2009 (ESSPIN 2009) (Jigawa, Kaduna, Kano, Kwara and Lagos). This research suggested that SBMCs were not functioning well: there was a lack of clarity and understanding over the SBMC's role and responsibilities; they lacked the financial resources to support schools in the ways that LGEAs often expected them to; community members were sometimes excluded by local elites; and there was little participation by women and children despite guidelines requiring their inclusion.

In this context, SBMCs were starting from a low base and with substantial sociocultural barriers to be overcome to reach functionality and inclusive participation. ESSPIN has aimed to improve community involvement in schools through functioning SBMCs and increased women's and children's participation, with a number of interventions under its Output stream 4 (see Annex C). By 2013/14 this type of support had been extended to 17% of Kaduna's primary schools. It is hoped that SBMC functionality and inclusiveness will have improved substantially, at least in this group of schools, although the coverage of the programme may not yet be large enough for state-wide averages to have risen measurably.

By the time of the first round of the Composite Survey, our sample suggests that nearly all schools in Kaduna had SBMCs, and by CS2, all of the schools in the sample had SBMCs (Table 18). This does not mean that all the SBMCs are functional or inclusive, however – they may not have received ESSPIN Output stream 4 support. In Kaduna, women's and children's participation training began only in 2012/13, and within a minority of schools, so we cannot expect high performance on these indicators in CS1 or any substantial impact by the time of CS2. The following sections use criteria and standards defined by the ESSPIN logframe to examine SBMC functionality and the extent to which SBMCs are inclusive of women and children.

Table 18. Kaduna: Sample size of schools with SBMCs

	CS1	CS2
Schools sampled in Kaduna	105	140
Schools with SBMCs sampled in Kaduna	99	140

3.5.1 SBMC functionality

There are nine criteria used to assess SBMC functionality, of which five must be fulfilled to meet the logframe standard (Box 8). In Kaduna there was little difference in the state average between CS1 and CS2, apart from a reduction in the number of SBMCs which also have a children's committee (criterion 7) (Table 19). There was no significant change in the proportion of SBMCs meeting the overall standard for functionality. Most of the criteria for SBMC functionality rely on the ability to provide written or photographic evidence, or at least oral recollection of a specific event. Consequently, the criteria may reflect the quality of record keeping of the SBMC more than the particular aspects of functionality that they aim to measure. Also, as explained above, improving the functionality of SBMCs may be very challenging given the structural cultural barriers to overcome.

There was a significant increase in two additional criteria related to the inclusiveness of SBMCs between 2012 and 2014 which were not included in the CS1 report. These were whether the SBMC did anything to support commonly excluded groups, and whether it raised issues of children's exclusion from school with the community, LGEA or state government.

Box 8. Logframe standard for SBMC functionality

The school must meet at least five of the nine criteria listed below in order to meet the SBMC functionality standard for the current school year:[3]

1) Two or more SBMC meetings have taken place since the start of the current school year (written evidence);
2) SBMC conducted awareness-raising activities (written or oral evidence);
3) SBMC took steps to address exclusion (written or oral evidence);
4) SBMC networked with Community-Based Organisations (CBOs), traditional or religious institutions, or other SBMCs (written or physical evidence);
5) SBMC interacted with local government education authorities on education service delivery issues (written or physical evidence);
6) An SBMC women's committee exists (written or physical evidence);
7) An SBMC children's committee exists (written or physical evidence);
8) SBMC contributed resources for the school (written or physical evidence); and
9) SBMC chair visited the school at least three times from the start of the current school year (written evidence).

Table 19. Kaduna: SBMC functionality in CS1 and CS2

	CS1	CS2	
(1) 2+ meetings this school year	18.9	6.5	
(2) Conducted awareness-raising	33.6	20.1	
(3) Addressed exclusion	27.5	17	
(4) Networked with CBOs/institutions/other SBMCs	21.4	36.5	
(5) Interacted with LGEA	27.6	24.7	
(6) Has women's committee	17.3	31.5	
(7) Has children's committee	27.5	12.4	-
(8) Contributed resources for school	48.1	49.2	
(9) Chair visited school 3+ times	31	13.1	
Schools meeting functioning SBMC standard	28	21.5	
Number of SBMC functionality criteria met (/9)	2.6	2.1	
Additional criteria			
Action for commonly excluded groups	2.1	11	+
Raised issue of children's exclusion	3.9	17	+

Note: + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

Looking at the difference between ESSPIN and non-ESSPIN schools, in 2014 Kaduna's ESSPIN schools performed more effectively on many of the SBMC functionality criteria (Table 20). ESSPIN schools' SBMCs were significantly more likely to have held two or more meetings this year, taken steps to address exclusion, and had the SBMC chair visit the school at least three times this year. The average number of criteria met was higher in ESSPIN schools – at 3.4 out of 9 – compared with non-ESSPIN schools (1.8 out of 9). Schools with ESSPIN interventions also had better results than the control schools in the two additional criteria of inclusiveness, with the difference significant

[3] A slightly different standard with 10 criteria was used in CS1. The new standard with nine criteria was applied to both the CS1 and CS2 data.

for SBMCs raising issues of children's exclusion. These results suggest that the ESSPIN schools have benefited from the Output 4 interventions as expected, so increasing the structured role of the community in school management.

Table 20. Kaduna: SBMC functionality in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) 2+ meetings this school year	1.5	27.7	+
(2) Conducted awareness-raising	15.5	39.8	
(3) Addressed exclusion	10.2	45	+
(4) Networked	30.9	60.1	
(5) Interacted with LGEA	25.3	21.8	
(6) Has women's committee	30.9	34.1	
(7) Has children's committee	11.1	18.1	
(8) Contributed resources for school	46	62.6	
(9) Chair visited school 3+ times	8.6	32.1	+
Standard G: functioning SBMC	18.4	34.8	
Number of SBMC functionality criteria met (/9)	1.8	3.4	+
Additional criteria			
Action for commonly excluded groups	7.6	25.5	
Raised issue of children's exclusion	10.2	45	+

Note: + indicates a significant difference from the results in non-ESSPIN schools.

Although the overall level of SBMC functionality did not significantly increase between the surveys, we are interested in establishing whether SBMCs which received more ESSPIN support improved faster than SBMCs in other schools. The comparison of means method does not find a significant effect of the ESSPIN activities on SBMC functionality between 2012 and 2014 (Table 21), suggesting that the change in functionality was the same whether or not the SBMC had benefited from ESSPIN support in recent years.

Table 21. Kaduna: SBMC functionality difference in differences (comparison of means)

Number criteria met (/9)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	2.2	3.6
CS2	1.8	3.4
Difference	-0.4	-0.1

Note: * indicates a significantly different difference than that in non-ESSPIN schools ($p < .05$).

3.5.2 Women's inclusiveness

The 2009 study of SBMCs found that community members were excluded from the process by local elites, and as such SBMCs were little known about and lacked a link to the community. In addition, the requirement (as stated in the guidelines) for participation by women and students was often ignored where this was felt inappropriate in the local culture. In this section and the following section, we examine the extent to which SBMCs were inclusive of women's and children's concerns in 2012 and 2014. We measure SBMC women's inclusiveness using four criteria (Box 6). The support to improve the participation of women and children, which included establishing sub-committees, was rolled out after initial SBMC training and mentoring. In Kaduna this was not until

2012/13, and only in a subset of schools (see Annex C), so we would not expect to see much impact on women and children's participation at the time of the CS2.

In fact in Kaduna there was a significant improvement between CS1 and CS2 in the number of SBMCs where at least one women's committee meeting took place (Table 22). The number of all schools meeting the overall standard for SBMC women's inclusiveness remained the same in CS2 as in CS1.

Box 9. Logframe standard for SBMC women's inclusiveness

The school must meet at least three of the four criteria listed below in order to meet the SBMC women's inclusiveness standard for the last school year:

1) At least one woman attended two or more SBMC meetings (written evidence);
2) Female member of SBMC raised at least one issue at SBMC meetings (written evidence or oral evidence from a female member of the SBMC);
3) At least one issue raised by a female member at an SBMC meeting led to action (written, physical or oral evidence from a female member of the SBMC); and
4) At least one SBMC women's committee meeting took place.[4]

Table 22. Kaduna: SBMC's women's inclusiveness in CS1 and CS2

	CS1	CS2	
(1) At least one woman attended 2+ meetings (%)	9.2	3.2	
(2) Female member raised an issue (%)	14.6	29	
(3) Issue raised by female member led to action (%)	31.1	8.5	
(4) Women's committee met (%)	1.6	24	+
Number of criteria met	0.5	0.6	
Meets standard (3/4 criteria)	5.9	4	

Note: Six schools that did not have SBMCs at all in CS1 are excluded from the analysis; + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

In the CS2 results, SBMCs expected to have improved due to ESSPIN are more likely to have had a woman attend at least two meetings than schools not expected to have improved (Table 23). The number of SBMCs meeting the overall standard for women's inclusiveness was also found to be higher and significant in ESSPIN schools than in non-intervention schools, at 14% of ESSPIN SBMCs compared with only 2% in other schools.

Table 23. Kaduna: SBMC women's inclusiveness in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) At least one woman attended 2+ meetings (%)	0.4	15.2	+
(2) Female member raised an issue (%)	27.7	34.7	
(3) Issue raised by female member led to action (%)	5.1	22.1	
(4) Women's committee met (%)	24.4	22.2	
Number of criteria met	0.5	0.9	
Meets standard (3/4 criteria)	1.6	14.1	+

Note: + indicates a significant difference from the results in non-ESSPIN schools.

[4] This criterion has been slightly altered since CS1, where it also required the women's committee to have a female leader.

The difference in differences analysis finds no significant difference in improvement of SBMC women's inclusiveness for ESSPIN schools, implying that the SBMC women's inclusiveness did not improve any faster in schools expected to benefit from ESSPIN over the period than other schools (Table 24).

Table 24. Kaduna: SBMC women's inclusiveness difference in differences (comparison of means)

Number criteria met (/4)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	0.5	0.4
CS2	0.5	0.9
Difference	0	0.5

Note: Six schools that did not have SBMCs at all in CS1 are excluded from the analysis; * indicates a significantly different difference than in non-ESSPIN schools ($p < .05$).

3.5.3 Children's inclusiveness

Earlier qualitative research (ESSPIN, 2009) found that many SBMCs did not allow the participation of children, and that where they had student members, they were not always able to be invited or may not have been comfortable voicing opinions in meetings. In this section, we examine whether SBMCs have improved in the extent to which they are inclusive of children, in accordance with guidelines on how they are supposed to operate. There are four criteria within the standard on SBMC children's inclusiveness.

As mentioned above, the intervention support for children's participation was first rolled out only to a selection of schools in 2012/13, and ESSPIN Output 4 interventions had reached only a minority of all schools (17%) by CS2. It may be unsurprising therefore that in Kaduna only 1.3% of schools' SBMCs met the overall standard in CS2 (Table 25). There was little change between CS1 and CS2 in the proportion of schools meeting the standard, and although each of the individual criteria improved, the changes were not significant. There was, however, a significant improvement in the average number of criteria met, now at 0.4 (out of 4) compared with 0.1 in CS1.

Box 10. Logframe standard for SBMC children's inclusiveness

The school must meet at least three of the four criteria listed below in order to meet the SBMC's children's inclusiveness standard for the current school year:

1) At least one child attended two or more SBMC meetings (written evidence);
2) A child member of SBMC raised at least one issue at SBMC meetings (written evidence or oral evidence from child member of SBMC);
3) At least one issue raised by a child member at an SBMC meeting led to action (written, physical or oral evidence from child member of SBMC); and
4) At least one SBMC children's committee meeting took place and the committee has a trained facilitator.[5]

[5] In CS1 this criterion required written evidence in the form of minutes of at least one children's committee meeting held in the past school year. This requirement was dropped for CS2 as it was considered unlikely that children's committees would keep good minutes, and that a failure to keep minutes does not mean the committee is not functioning.

Table 25. Kaduna: SBMC children's inclusiveness in CS1 and CS2

	CS1	CS2	
(1) Child attended 2+ meetings (%)	1.6	1.8	
(2) Child raised an issue (%)	5.3	18.8	
(3) Issue raised by child led to action (%)	1.3	2.8	
(4) Children's committee met and it has a trained facilitator (%)	0.4	15.1	
Number of criteria met	0.1	0.4	+
Meets standard (3/4 criteria) (%)	0.5	1.3	

Note: Six schools that did not have SBMCs at all in CS1 are excluded from the analysis; + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

Looking at children's inclusiveness in CS2, the ESSPIN schools performed better than non-ESSPIN schools on all the indicators (Table 26), and significantly so in terms of children attending SBMC meetings. Overall, 7% of ESSPIN schools met the standard for SBMC children's inclusiveness, while none of the non-ESSPIN schools did.

The difference in differences analysis confirms that all schools improved, and suggests that ESSPIN schools improved at a similar rate to non-ESSPIN schools between 2012 and 2014 (Table 27). Regression analysis leads to the same finding, as shown in Annex D.

Table 26. Kaduna: SBMC children's inclusiveness in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) Child attended 2+ meetings (%)	0	9.1	+
(2) Child raised an issue (%)	14.8	35.5	
(3) Issue raised by child led to action (%)	1.9	6.4	
(4) Children's committee met and it has a trained facilitator (%)	14.4	17.6	
Number of criteria met	0.3	0.7	
Meets standard (3/4 criteria) (%)	0	6.9	+

Note: + indicates a significant difference from the results in non-ESSPIN schools.

Table 27. Kaduna: Difference in differences in SBMC children's inclusiveness (comparison of means)

Number of criteria fulfilled (/4)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	0	0.2
CS2	0.3	0.7
Difference	0.3	0.5

Note: Six schools that did not have SBMCs at all in CS1 are excluded from the analysis; * indicates a significantly different difference than that in non-ESSPIN schools ($p < .05$).

Overall, Kaduna saw an improvement in the extent to which SBMCs were inclusive of children and no clear change in the extent to which they were inclusive of women. Much higher proportions of ESSPIN than non-ESSPIN schools met the inclusiveness standards in 2014. But most schools remain well below the standards they are expected to meet. When we looked specifically at the rate of change between CS1 and CS2, we found no clear difference between ESSPIN and non-ESSPIN schools, in terms of either women's or children's inclusion.

3.6 School quality

Overall school quality is measured as a combination of the standards on teacher competence, headteacher effectiveness, school development planning, and SBMC functionality. A high-quality school is defined as one that meets the teacher competence standard and at least two of the other standards (Box 11). We find that 1% of schools met the overall school quality standard in CS1, and 2.8% in CS2. However, this change is not statistically significant (Table 28). We also use a 'quality score' indicator which is an average of the continuous indicators developed in the previous sections for teacher competence, headteacher effectiveness, school development planning, and SBMC functionality. There was a small decrease in this quality score from 38% in CS1 to 36% in CS2, though this was again not significant.

Box 11. Logframe standard for school quality

The school must meet at least three of the four output standards listed below in order to meet the school quality outcome standard, with teacher competence having to be one of those three:

1) Teacher competence standard (more than 50% of sampled teachers are competent);
2) Headteacher effectiveness standard;
3) School development planning effectiveness standard; and
4) SBMC functionality standard.

As the teacher competence standard has changed between CS1 and CS2 – with teachers required to score a minimum of 50% in both English and mathematics tests to pass the competence standards – we report both 'CS1' and stricter 'CS2' versions of the overall quality standard.

Table 28. Kaduna: School quality in CS1 and CS2

	CS1	CS2	
Meets three or four standards (CS1 version) (%)	1.0	2.8	
Quality score (CS1 version)	38.3	36.1	

Within CS2, the results suggest that schools which were expected to have improved due to ESSPIN were of higher quality than other schools, for example 15% of ESSPIN schools met the school quality standard (as defined in CS1) compared with none of the non-ESSPIN schools. The differences between the ESSPIN and non-ESSPIN results are not significant, which reflects the high degree of variation within each group of schools in our survey sample in Kaduna (Table 29).

Table 29. Kaduna: School quality in CS2, by intervention group

	(i) Non-ESSPIN	(ii) ESSPIN	
Meets three or four standards (CS1 version)	0	14.6	
Meets three or four standards (CS2 version)	0	1.5	
Quality score (CS1 version)	34.3	43.5	
Quality score (CS2 version)	33.5	41.9	

Note: The CS2 version of the quality score and school quality standard reflect the 'strict' version of the teacher competence standard, where teachers are required to pass literacy and numeracy tests as well as fulfilling other criteria; + indicates a significant difference from the results in non-ESSPIN schools.

The difference in differences analysis of means find no conclusive evidence that ESSPIN schools have improved more than schools where no improvement is expected in terms of school quality

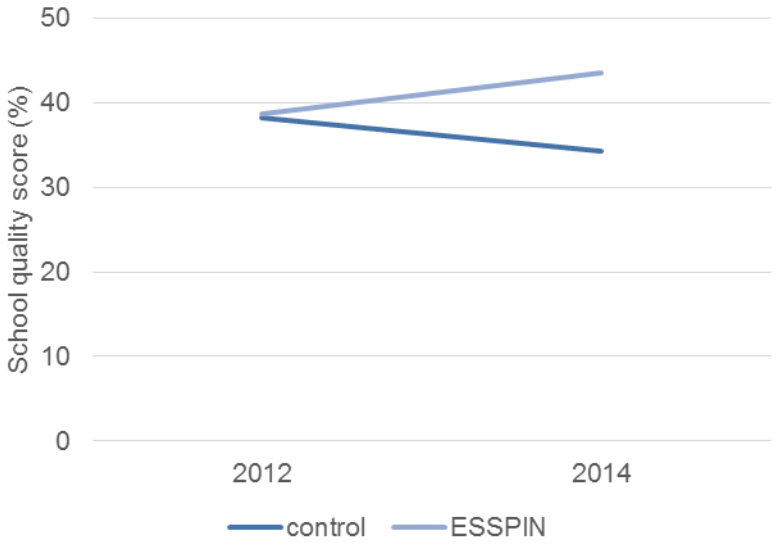
(Table 30). However, the direction of the results is that ESSPIN schools have improved relative to other schools, as can be seen in Figure 4. Although the difference in the change is not significant, the direction provides some limited support for the hypothesis that the SIP helped to improve quality during 2012 to 2014. The regression method, shown in Annex D, does not find strong evidence of any faster improvement in school quality due to ESSPIN.

Table 30. Kaduna: School quality difference in differences (comparison of means)

School quality score	(i) Non-ESSPIN	(ii) ESSPIN
CS1	38.2	38.7
CS2	34.3	43.5
Difference	-3.9	4.8

Note: * indicates a significantly different difference than that in non-ESSPIN schools ($p < .05$).

Figure 4. Kaduna: School quality in 2012 and 2014, in control and ESSPIN schools



3.7 Pupil learning achievement in English literacy and numeracy

The learning achievement indicators are included as indicators of the impact of the ESSPIN programme. Pupils were tested in grades 2 and 4, in literacy and numeracy. This section begins by following the same analysis conducted for other indicators: looking at the change in the state average between CS1 and CS2, the difference between ESSPIN and non-ESSPIN schools in 2014, and the change in results between 2012 and 2014 for pupils in schools expected to improve in this time compared with those not expected to improve. It then moves on to a more detailed look at the breakdown of pupil results. Here we look at how pupil scores were distributed in the 2014 tests, split between ESSPIN and non-ESSPIN schools. We also look at the average test scores on sub-scales of the tests, such as grasp-of-number concepts or addition and subtraction, and how this varies over time for the state average and for pupils from different types of schools.

3.7.1 Main analysis

The average test scores in Kaduna range from 21% in grade 4 literacy to 34% in grade 2 numeracy. In 2014 scores are generally higher in grade 2 than grade 4, and higher in numeracy than literacy. Between 2012 and 2014 there was a drop in the average pupil test scores in Kaduna across the four tests: literacy and numeracy in grades 2 and 4; and the drop was significant in all

subjects, except grade 2 literacy (Table 31). The logframe indicators, although tightly defined, give an indication of the proportion of students with the ability to read/answer numeracy questions at the appropriate level for their grade. The results for Kaduna in these indicators are low, with less than 6% of pupils meeting the level in all four tests and less than 1% in grade 4 tests of literacy and numeracy. The numeracy logframe indicators also saw a significant decline between CS1 and CS2 in Kaduna.

Table 31. Kaduna: Test scores and proportion of children reaching logframe indicator in CS1 and CS2

	Test	CS1	CS2	
Test score (%)	L2	31.8	26.8	
	L4	31.5	20.6	-
	N2	53.5	33.7	-
	N4	41.2	28.7	-
Logframe indicator (%)	L2	2.2	1.2	
	L4	1.4	0.8	
	N2	15.7	5.4	-
	N4	10.3	0.4	-

Note: + = significant improvement between 2012 and 2014; - = significant worsening between 2012 and 2014 (using a t-test; $p < .05$).

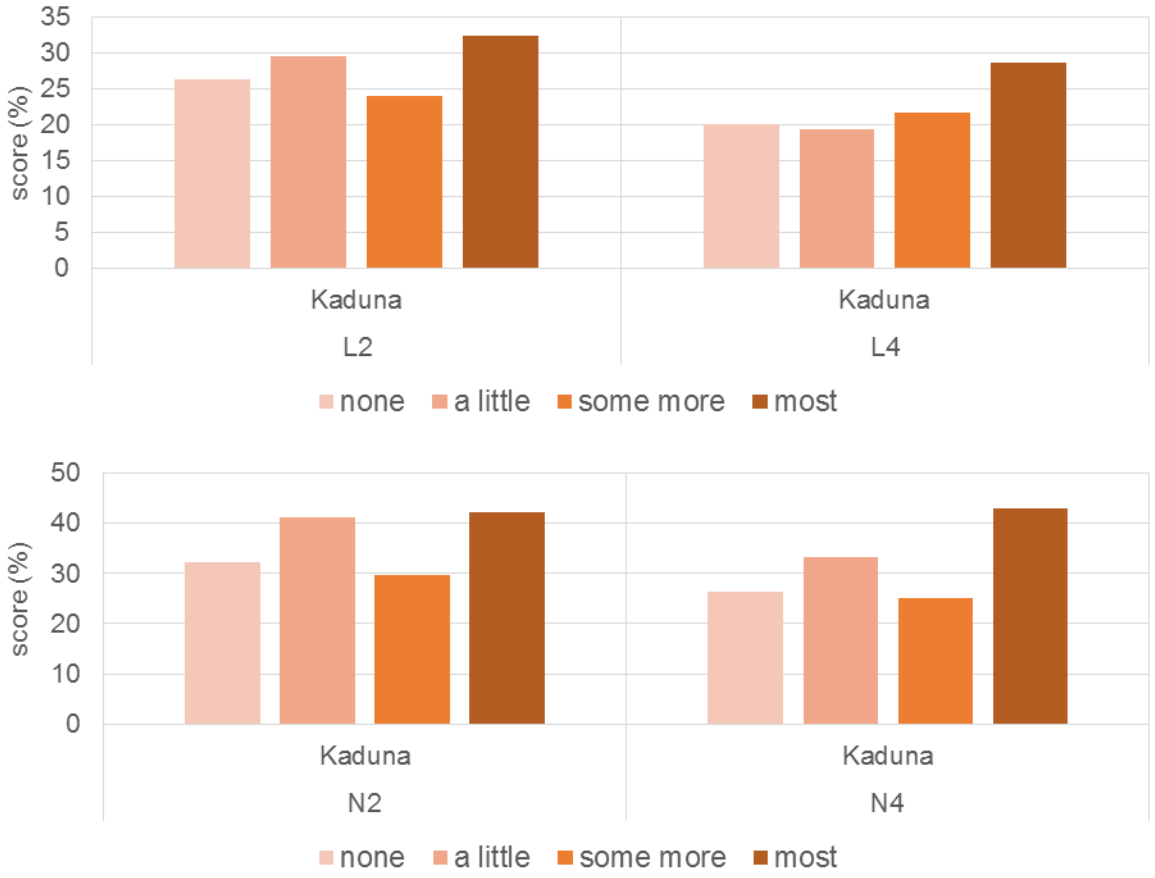
Focusing on the CS2 data, pupils who had spent time in schools whose quality was expected to have improved as a result of ESSPIN had higher test scores than those in control schools, although the difference was not significant (Table 32). There was also no difference found between the ESSPIN and non-ESSPIN schools in terms of meeting the logframe indicators. The lack of significance here may be explained by the variation between schools and pupils in Kaduna. In addition to the general variation mentioned earlier, the ESSPIN treatment group is heterogeneous, with some schools having only had one or two years of intervention. When we look at the results for each of these separate groups, we can see that schools receiving some (as opposed to the most) intervention had lower test scores, and these will have brought down the average results for pupils in ESSPIN schools (Figure 5).

Table 32. Kaduna: Test scores and proportion of children reaching logframe indicator in CS2, by intervention group

	Test	(i) Non-ESSPIN	(ii) ESSPIN	
Test score (%)	L2	26.3	27.9	
	L4	20	21.6	
	N2	32.4	36.9	
	N4	26.5	32.4	
Logframe indicator (%)	L2	1.5	0.7	
	L4	0.8	0.6	
	N2	4.5	7.6	
	N4	0.6	0.2	

Note: + indicates a significant difference from the results in non-ESSPIN schools.

Figure 5. Kaduna: Test scores by ESSPIN intervention group



Turning to the question whether the ESSPIN activities had an impact on change in pupil test scores between 2012 and 2014, we find no statistical difference between the rate of decline in the schools' scores. The difference in differences analysis confirms that pupils' test scores were generally higher in ESSPIN schools than non-ESSPIN schools (Table 33), and that scores fell between CS1 and CS2 in both types of schools. There is some suggestion that ESSPIN schools' scores fell more than control schools in the tests (except for grade 4 numeracy), although this was not statistically significant. Figure 6 shows these results graphically. Regression analysis, reported in Annex D, also finds no significant impact of the ESSPIN intervention on improvement in test scores.

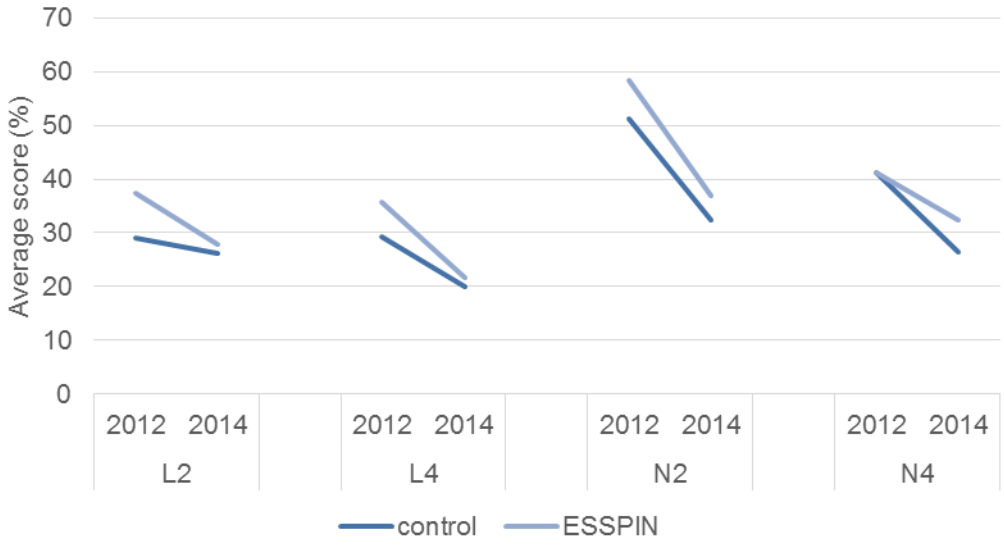
Table 33. Kaduna: Pupil test score difference in differences (comparison of means)

Pupil test score		(i) Non-ESSPIN	(ii) ESSPIN
L2	CS1	29	37.5
	CS2	26.3	27.9
	Difference	-2.7	-9.6
L4	CS1	29.3	35.6
	CS2	20	21.6
	Difference	-9.3	-14
N2	CS1	51.1	58.3
	CS2	32.4	36.9
	Difference	-18.8	-21.4
N4	CS1	41.2	41.1

	CS2	26.5	32.4
	Difference	-14.7	-8.7

Note: * indicates a significantly different difference than that in non-ESSPIN schools (p < .05).

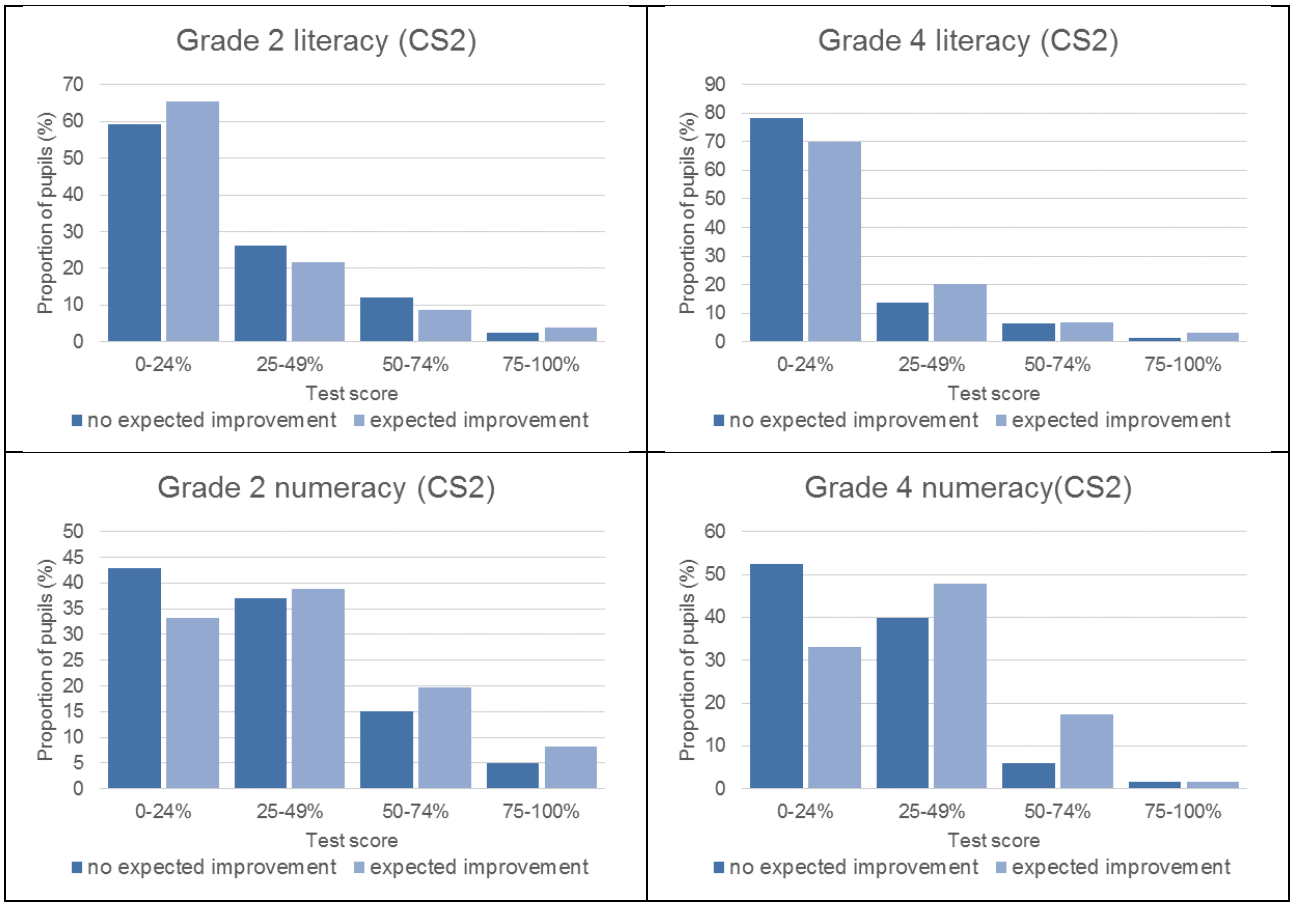
Figure 6. Kaduna: Pupil test scores in ESSPIN and control schools, in 2012 and 2014



3.7.2 Distribution of test scores and sub-scale scores

In Kaduna, in both schools expecting improvement and those not, very low proportions of pupils scored over 50% in all four subjects in CS2 (Figure 7). In literacy, over 59% of pupils scored 0–24% in both types of schools. In all tests apart from grade 2 literacy, pupils from schools expecting an improvement were more likely to score in the highest band (75–100%) than pupils from schools not expecting an improvement from ESSPIN. The numeracy charts in particular show the results that we would hope to see: there is a shift to the right, such that pupils in the schools expecting improvement are more likely to fall into the higher score bands.

Figure 7. Kaduna: Distribution of pupil test scores in 2014



The average pupil test scores in Kaduna have fallen between CS1 and CS2 in all sub-scale areas of the tests (Figure 8). For example, the grade 2 pupils scored 19% on average on grade 2 numeracy items in CS2, compared with an average of 42% in CS1.

Across many of the sub-scales of the pupil tests, pupils in the ESSPIN schools scored higher on average than pupils in the non-ESSPIN schools (Figure 9). Generally, this difference was quite marginal but was statistically significant for grade 4 pupils at grade 4 level numeracy items, in which ESSPIN school pupils scored 12% on average but non-ESSPIN school pupils 6%.

Figure 8. Kaduna: Average scores in test sub-scales, CS1 and CS2

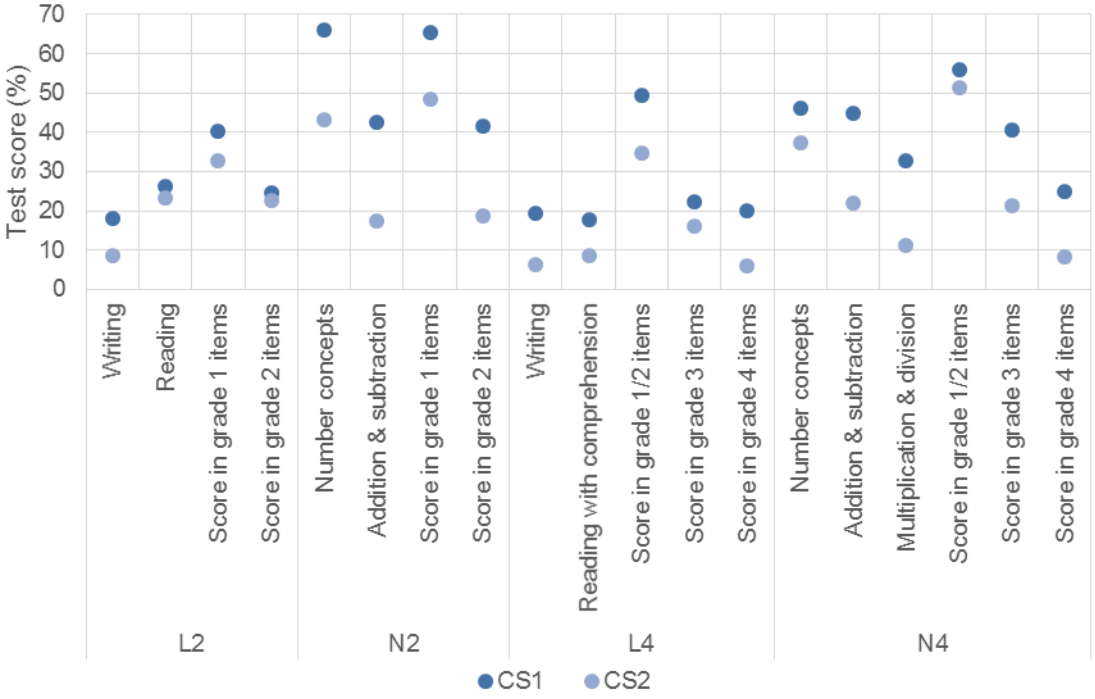
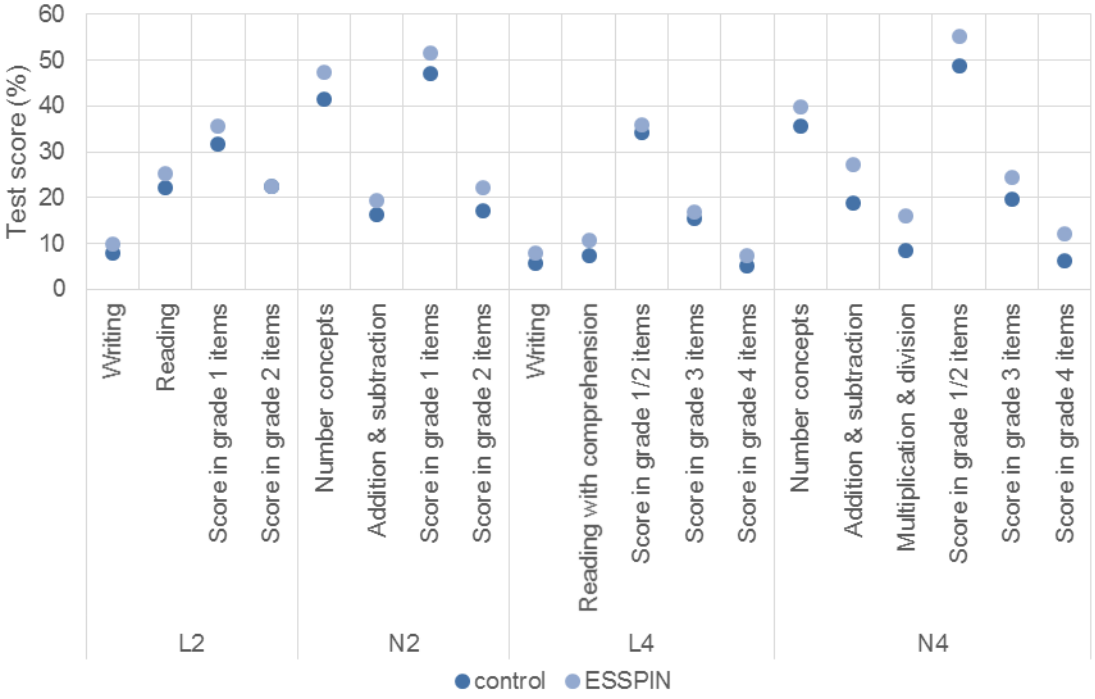


Figure 9. Kaduna: Average scores in test sub-scales in CS2, ESSPIN and non-ESSPIN



4 Conclusion and implications of survey findings for Kaduna's ESSPIN programme

This report has looked at a set of indicators of programme outputs, outcomes and impact, defined under ESSPIN's theory of change. The results across the indicators for Kaduna follow a similar pattern:

- For schools in Kaduna as a whole, some measures of headteacher effectiveness, SBMC action on inclusion, and the extent to which SBMCs encourage children's participation, have improved between 2012 and 2014. However, teacher competence, overall school inclusiveness, and children's test results have fallen significantly.

- Schools that received ESSPIN intervention are better in terms of teacher competence, inclusiveness, SBMC functionality, and school development planning (although the differences are not always large and in some cases not statistically significant). They are also better in terms of overall school quality, and appear to have improved between 2012 and 2014, while other schools in the state have become worse.

- Pupil test results in literacy and numeracy remain low and do not differ significantly between ESSPIN and control schools.

These findings are affected by the high variability in Kaduna's schools, which makes it hard for us to provide precise estimates of the measures of school functioning and learning outcomes that we are interested in. As a result, sometimes the findings are in the expected direction but not statistically significant. However, this does not explain the results which were in the 'wrong' or unexpected direction, such as the fall over time in pupil test results.

State averages for indicators measured in the Composite Surveys did not see the improvement hoped for between 2012 and 2014 more generally in Kaduna. Around a quarter of all schools had benefited from at least some ESSPIN targeted support by the time of the CS2 survey, and it is expected that these schools would improve more than others and bring up the average. There were also some state-level activities which might be expected to improve performance in all schools. The absence of improvement implies that schools' capability to provide improved learning outcomes – through better teaching, learning, and community engagement and inclusiveness – remained stagnant, despite significant effort to bring up the overall performance.

Within ESSPIN schools, there is evidence of better teaching and overall quality, as well as inclusiveness and SBMC functionality. But this has not yet translated into better learning outcomes. This is despite the fact that we focus on schools that have had ESSPIN intervention for more than one year, and that schools receiving more ESSPIN intervention tend to be urban and larger than control schools, characteristics which are usually associated with higher test results.

What could explain this pattern? One possibility is that our measures of school functioning, many of which are oriented towards measuring specific processes or behaviours, do not capture the aspects that are truly important for children's learning outcomes. There may be bottlenecks or individual factors that can prevent learning from improving even where other factors are favourable.

Teachers' subject knowledge is likely to be one such bottleneck, preventing improved school functioning from leading to improved learning. As section 3.1 shows, many teachers struggle with basic literacy skills, especially writing in English and foundational concepts in English literacy. Most teachers can complete grade 1 level items in English and mathematics satisfactorily, but for items

pitched at primary grade 5, they could answer correctly only around 40% of mathematics items and 20% of English items. Although ESSPIN aims to improve subject knowledge through the provision of lesson plans, it is not clear that this is sufficient to overcome the severe constraints in teachers' subject knowledge, especially given that textbooks are written in English.

As discussed in section 1.2, Kaduna has also experienced an ongoing level of conflict and violence which would very likely have hindered educational improvement in the state. These threats and pressures on the communities in Kaduna are likely to have an ongoing negative impact on education and on many of the indicators measured here. SBMCs might be less likely to meet and function properly, and the stress on teachers and heads might mean that they are less effective. These conditions, along with high absenteeism for teachers and pupils, will have had a notable effect on learning outcomes.

Another contextual feature in Kaduna is a large increase in pupil enrolment across all schools, which led PTRs to increase by more than 50% across the state (see Annex A). More investigation is needed to understand whether this huge increase in enrolment represents a real increase in school participation, and whether ESSPIN interventions aiming to improve access have played a role in it. In the schools categorised as maximum ESSPIN intervention, the PTR almost doubled. With larger class sizes and stagnant inputs, it may not be surprising that teacher competence was not able to improve, and that pupil learning outcomes fell.

References

ESSPIN (2009) School Based Management Committees in Policy and Practice, Volume 1: Research Synthesis Report. Report number ESSPIN 404

ESSPIN (2013a) Overall findings and technical report of ESSPIN Composite Survey 1 (2012). Report number ESSPIN 060

ESSPIN (2013b) Kaduna State report of ESSPIN Composite Survey 1 (2012). Report number KD201

ESSPIN (2013c) Extension of the Education Sector Support Programme in Nigeria, August 2014–January 2017 [business case for the UK Department for International Development]

ESSPIN (2014) Education, Conflict and Violence Research. Report number ESSPIN 428

Megill, D.J. (2014) Final Sample Design and Weighting Procedures for the ESSPIN Second Composite Survey (CS2) in Nigeria. Unpublished note written for ESSPIN and Oxford Policy Management

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010) Introducing ACLED-Armed Conflict Location and Event Data. *Journal of Peace Research*, vol. 47, no. 5, 1–10. ACLED Version 5 (1997-2014) data accessed 28 January 2015 www.acleddata.com/

RTI International (2014) Nigeria reading and access research and activity: adaptation of Education Sector Support Programme in Nigeria's teacher capacity development strategy. Document produced for the United States Agency on International Development

Annex A School characteristics

The table below sets out summary statistics for Kaduna's schools, split by categories according to the level of Output stream 3 intervention (none, minimum, medium, maximum). The data comes from the annual school census from 2009/10 and 2013/14.

Kaduna's schools by level of ESSPIN intervention	None	Min.	Med. (1)	Med. (2)	Max.	
	0	2c	2b	2a	1	
Distance from local government authority headquarters	36.6	24.1	18.5	20.6	37	-
Age of the school in 2014	21.1	26.6	27.6	26.8	26.9	+
Urban (%)	6.5	12.8	32.4	18	12.3	+
Nomadic (%)	5.9	5.5	4.1	3.5	2.5	-
Islamic (%)	1.1	0	0.6	0	0	-
Double shift (%)	1.6	5.9	10.6	6.3	3.7	+
Had parent–teacher association in 2009/10 (%)	95.8	97.1	92.3	94.8	93.1	-
Had SBMC in 2009/10 (%)	83.8	88.1	81.4	87.9	84.8	
PTR in 2009/10	38.1	29.4	30.5	34	22.9	-
PTR in 2013/14	49.6	38.1	38	38.9	35.5	-
Change in PTR (%)	68.9	66.2	57.4	59.5	93.2	
Number of classrooms in 2009/10	4.9	4.8	5.4	5.1	4.9	
Number of teachers in 2009/10	7.2	13.8	16.2	15.3	17.2	+
Primary enrolment in 2009/10	213.4	332.2	424.3	341.8	285.2	+
Change in enrolment 2009/10–2013/14 (%)	40.7	50.7	17.2	46.7	76	+
% of teachers with academic diploma/degree	44.5	48.4	50.8	42.9	42.9	
% of teachers with PGDE, BEd or MEd	4.6	3.9	4.3	4.7	3.9	
% of teachers with NCE, Grade II or equivalent	78.8	79.1	81.8	79.7	78.6	
School has a power source (grid/other)	3	6.4	3.1	6.3	5.6	+
% of classrooms with enough seating	32.3	35.6	29.5	41.9	20.5	
% of classrooms with a good blackboard	43.4	46.3	42.3	52.4	36.7	
% of classrooms in good condition/minor repairs	53	55.2	58.4	61.5	47.4	
School has at least one toilet (%)	15.9	23.8	32.7	22.9	21	+
Number of schools	3068	293	170	431	163	

Source: Annual School Census, 2009/10 and 2013/14; +/- indicates a significant positive/negative coefficient in a linear or logit regression of years of full ESSPIN intervention on the variable of interest.

Annex B ESSPIN Output stream 3 interventions

The table below shows the ESSPIN Output stream 3 interventions delivered to date in Kaduna State. In order to make the variation in interventions across and within states manageable for analysis, each combination of interventions was categorised as none, minimum, medium, or maximum, according to the number of years of continuous intervention.

Expected impact	Number of schools	2009/10			2010/11			2011/12			CS1	2012/13			2013/14			CS2
		L	T	SV	L	T	SV	L	T	SV		L	T	SV	L	T	SV	
None	3,068																	
Minimum	293														6	3	9	
Medium (1)	240											6	3	9	6	3	9	
Medium (2)	317							6	3	9		6	3	9	3		9	
Maximum	165	5*	5*	9*	10*	5*	9*	6	3	9		6	3	9	3		9	

Note: L = days of leadership training; T = days of teaching training; SV = school visits; * = pilot.

Annex C ESSPIN Output 4 interventions

The table below shows the days of Output 4 intervention in Kaduna under different headings: SBMC training; women and children participation training; and mentoring visits.

Level of Output stream 3 intervention	2010/11			2011/12			CS1	2012/13			2013/14			CS2	De facto phase
	S	P	M	S	P	M		S	P	M	S	P	M		
None/min/med															control
Med				7		4		r		4		6	4*		pre-CS1
Max	7		4	r		4			6	4*			4*		pre-CS1

Note: S = SBMC training; P = women and children participation training; M = mentoring visits; r = one-day refresher; mentoring visits were by civil society-government partnership teams, except those marked with an asterisk, which were by social mobilisation officers.

Annex D Difference in differences analysis using regression

This annex presents the results of regression analysis of difference in differences for the indicators above, to compare change over time between subgroups.

The analysis in Chapter 3 above presents the difference in differences of indicator means for sub-categories of groups.

Sometimes we want to use all of the available information and compare schools which have had more or less ESSPIN intervention – a continuous scale – rather than dividing them into some or none. In this case we can use regression analysis – a statistical process for estimating relationships among variables. We model the outcome indicator as depending on time (the round of the survey, CS2 versus CS1), the intensity of intervention, and a treatment effect, which is the interaction between time and intensity of intervention. The treatment effect tells us if an increase in the level of intervention increased the speed at which the outcome improved. Regression results are reported as a series of 'coefficients' – numbers representing the strength of the relationship with the outcome of interest.

Coefficient	Meaning of coefficient if positive and significant
Time (CS2 v. CS1)	the outcome improved over time
Intervention	the higher the level of intervention the more effective (regardless of change over time)
Treatment	the higher the level of intervention, the more or faster the outcome improved over time – this is our key indicator of success

Teacher competence

Here we use regression analysis to examine whether teachers who received more training from ESSPIN improved more, or faster, between 2012 and 2014 (Table 34). We model the outcome indicator (competence score) as depending upon time (the round of the survey) and the intensity of intervention. Intensity of intervention is measured in terms of the amount of teacher training delivered to the school. We also use an alternative intervention measure that adjusts for the length of time a teacher has been in his or her present school. A teacher who joined the school only in 2012, for example, cannot be expected to have benefited from ESSPIN training delivered in 2010 or 2011, and the intervention variable can be adjusted to reflect this.

The interaction effect between intervention and time, labelled 'treatment', if positive and significant, would provide evidence that schools with more ESSPIN intervention improved more rapidly between 2012 and 2014. The time effects in the non-adjusted results are negative and significant, suggesting that teachers' competence generally fell between CS1 and CS2. The intervention effect is not significant, so no difference between ESSPIN and non-ESSPIN schools' teachers can be seen. The treatment effects in Kaduna are also not significant, but they are positive and in the expected direction – with greater effects for teachers who received the training than those just in ESSPIN schools.

Table 34. Kaduna: Teacher competence difference in differences (regression)

Regression on competence scores (CS1 version)		Non-adjusted				Adjusted			
Intervention variable		school improvement		training		school improvement		training	
Time (CS2 v. CS1)	coefficient	-9.8	*	-9.5	*	-4.3		-5	
	SE	4.7		4.8		4.4		3.8	
Intervention	coefficient	-1.6		-1.2		-1.7		-2.7	
	SE	1.7		1.6		1.7		0	
Treatment	coefficient	2.6		5.2		0.4		3	
	SE	2		3.2		2.2		2.9	
	N	233		233		195		195	

Note: * indicates a significant coefficient (p < .05). Adjusted results are adjusted for the length of time a teacher has been in the current school – and therefore whether they would have benefited from the full ESSPIN training package.

Headteacher effectiveness

The regression results suggest there was no significant difference in the change in number of effectiveness criteria met by headteachers in schools benefiting from ESSPIN and schools which did not expect an improvement (Table 35). We also adjusted for the year that the headteacher was appointed to his or her current school and still found no evidence of a difference between the two groups.

Table 35. Kaduna: Headteacher effectiveness difference in differences (regression)

Regression on number of criteria met (out of 7)		intervention variable					
		School improvement		Training		Training (adjusted for start date)	
Time (CS2 v. CS1)	coefficient	-0.07		-0.17		-0.25	
	SE	0.27		0.29		0.28	
Intervention	coefficient	-0.06		-0.03		-0.03	
	SE	0.21		0.03		0.03	
Treatment	coefficient	-0.1		0.04		0.03	
	SE	0.38		0.06		0.06	
	N	205		205		172	

Note: * indicates a significant coefficient (p < .05).

School development planning

Regression analysis finds no significant difference in the change between ESSPIN and non-ESSPIN schools in terms of school development planning criteria (Table 36). This means that we cannot say that the ESSPIN interventions made schools more likely to improve in school development planning between 2012 and 2014.

Table 36. Kaduna: SDP effectiveness difference in differences (regression)

Regression on SDP effectiveness criteria fulfilled (/5)			
Time (CS2 v. CS1)	coefficient	-0.1	
	SE	0.13	
Intervention	coefficient	0.35	*
	SE	0.17	
Treatment	coefficient	0.07	
	SE	0.26	
	N	237	

Note: * indicates a significant coefficient (p < .05).

School inclusiveness

A regression on the inclusiveness score has been used for difference in difference analysis of schools in Kaduna depending on the level of ESSPIN intervention (Table 37). The results show that both groups of schools became less inclusive between 2012 and 2014 (as shown by the significant negative coefficient on time). As suggested by the insignificant treatment coefficient, there is no evidence to support the idea that ESSPIN schools improved (or declined) at a different rate to schools not expected to benefit from the programme.

Table 37. Kaduna: School inclusiveness difference in differences (regression)

Regression on inclusiveness score			
Time (CS2 v. CS1)	coefficient	-9.6	*
	SE	4.27	
Intervention	coefficient	1.29	
	SE	2.72	
Treatment	coefficient	-0.17	
	SE	2.55	
	N	233	

Note: * indicates a significant coefficient (p < .05).

SBMC functionality and inclusiveness

A regression to assess the impact of ESSPIN support on the change in SBMC functionality finds no significant treatment effect (Table 38). This suggests that the change in functionality was the same whether or not the SBMC benefited from ESSPIN support in recent years.

Table 38. Kaduna: SBMC functionality difference in differences (regression)

Regression on number of SBMC functionality criteria met			
time (CS2 v. CS1)	coefficient	-0.29	
	SE	0.46	
intervention	coefficient	1.15	
	SE	0.82	
treatment	coefficient	-0.05	
	SE	0.92	
	N	235	

Note: * indicates a significant coefficient (p < .05).

The regression on women's inclusiveness of SBMCs also finds that there was no significant treatment effect, so ESSPIN support did not lead to any faster improvement in women's inclusiveness between 2012 and 2014 (Table 39).

Table 39. Kaduna: SBMC women's inclusiveness difference in differences (regression)

Regression on number of women's inclusiveness criteria met			
Time (CS2 v. CS1)	coefficient	-0.01	
	SE	0.2	
Intervention	coefficient	0.02	
	SE	0.15	
Treatment	coefficient	0.28	
	SE	0.16	
	N	220	

Note: Six schools that did not have SBMCs at all in CS1 are excluded from the analysis; * indicates a significant coefficient (p < .05).

A regression analysis of difference in differences in SBMC children's inclusiveness finds that all schools improved between 2012 and 2014, according to the positive and significant time coefficient (Table 40). However, the treatment coefficient is not significant, suggesting all SBMCs became similarly more inclusive of children over the period, regardless of ESSPIN support received.

Table 40. Kaduna: Difference in differences in SBMC children's inclusiveness (regression)

Regression on number of children's inclusiveness criteria met			
Time (CS2 v. CS1)	coefficient	0.29	*
	SE	0.12	
Intervention	coefficient	0.1	
	SE	0.06	
Treatment	coefficient	0.11	
	SE	0.12	
	N	234	

Note: Six schools that did not have SBMCs at all in CS1 are excluded from the analysis; * indicates a significant coefficient (p < .05).

School quality

The regression method for difference in differences in school quality does not find a statistically significant treatment coefficient (Table 41). This suggests there was no significant difference in the change in school quality for schools expected to improve due to ESSPIN, compared with non-ESSPIN schools.

Table 41. Kaduna: School quality difference in differences (regression with continuous intervention variable)

Regression on school quality score			
Time (CS2 v. CS1)	coefficient	-2.55	
	SE	3.51	
Intervention	coefficient	4.15	
	SE	3.24	
Treatment	coefficient	1.26	
	SE	4.18	
	N	192	

Note: * indicates a significant coefficient (p < .05).

Pupil learning

In terms of pupil learning achievement, the regression analysis finds that in grade 4 literacy and grades 2 and 4 numeracy, average test scores fell between 2012 and 2014 in all schools (as shown by the negative and significant time coefficient in Table 42). The treatment coefficients are not significant for any of the four tests, suggesting that more ESSPIN support did not make schools' test scores any more likely to improve (or fall less) over the two years.

Table 42. Kaduna: Pupil test score difference in differences (regression)

Regression on pupil test score		L2		L4		N2		N4	
Time (CS2 v. CS1)	coefficient	-1.36		-6.42	*	-12.43	*	-13.9	*
	SE	3.89		2.59		4.71		5.58	
Intervention	coefficient	-1.06		-0.08		-0.14		-1.15	
	SE	1.97		1.37		2.15		1.47	
Treatment	coefficient	1.11		1		0.07		2.36	
	SE	1.57		0.95		1.85		1.48	
	N	244		243		244		243	

Note: * indicates a significant coefficient (p < .05).