Oxford Policy Management

ESSPIN Composite Survey 2

Jigawa 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 the findings for Jigawa 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 Jigawa 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 Jigawa the first 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. The programme was expanded by training a new cadre of support officers at the local government level, in addition to the existing state-level officers. By the time of the CS2 fieldwork, 48% of Jigawa's primary schools had benefited from at least one year of ESSPIN support. Schools which benefited from ESSPIN only in 2013/14 are treated as non-ESSPIN, since the support is not expected to have taken effect yet.

There have been large increases in enrolment in Jigawa since 2009, and although there have been a few instances of violent conflict recorded in the media, there have been recent anecdotal reports of violence or threats of violence, making it difficult for schools to work in parts of the state. 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 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 main findings from the Composite Surveys are as follows:

Teacher competence: In the second round of the survey, 19% of teachers who had received ESSPIN training met our standard for teacher competence, compared with only 3% of teachers in non-ESSPIN schools, a substantial and statistically significant difference. Teacher competence did not change significantly in Jigawa between CS1 and CS2: there was a significant positive improvement in the use of teaching aids and praise, but also there was a reduction in teachers' knowledge of benchmarks in the English or maths curriculum, from 57% in 2012 to 21% in 2014. Across Jigawa, teachers could answer fewer than one in 10 English questions at grade 5 level, and less than one-third of the maths questions at this level. When we look at change between the two surveys, there is no evidence that ESSPIN-trained teachers improved faster between 2012 and 2014 than teachers in non-ESSPIN schools.

Headteacher effectiveness: In Jigawa in 2014, 26% of headteachers who had received leadership training under ESSPIN met our effectiveness standard, compared to only 4% of those who had not benefited from training. There was no significant change in the proportion of headteachers who met the standard for effectiveness between CS1 and CS2. There was no evidence to suggest that the heads who had benefited from training had improved any faster (or worsened any less) than those without training in the period from 2012 to 2014.

School development planning: While only 6% of schools meet our school development planning standard in Jigawa in 2014, this was a large and significant increase on the 2012 figure (1%). Schools which have benefited from ESSPIN interventions are better at planning than those who have not benefited, with 24% meeting the standard compared with 1% of non-ESSPIN schools. ESSPIN schools also improved their school development planning faster than the control schools between CS1 and CS2.

School inclusiveness: In 2014 ESSPIN schools had a higher average inclusiveness score than non-ESSPIN schools. However, across Jigawa, average measures of school inclusiveness generally fell between CS1 and CS2: the proportion of Jigawa's schools meeting the overall standard was 5% in 2014, compared with 20% in 2012. Both ESSPIN and non-ESSPIN schools appear to have declined, but ESSPIN schools have declined more slowly (although this effect does not reach statistical significance).

SBMC functionality and inclusiveness: SBMCs which received ESSPIN training were significantly more functional than those without the training. They also have more participation from women and children, and were more likely to have raised the issue of children's exclusion with the school. Overall, there was no evidence of significant change in the average functionality of SBMCs between 2012 and 2014 in Jigawa, nor of change in the participation of women and children in SBMCs. However, some specific indicators improved, including the proportion that had raised issues of children's exclusion, and the proportion that had a women's committee which had met recently. SBMCs that received training in the relevant period appear to have improved between 2012 and 2014, while SBMCs in other schools may have worsened over time, in terms of both their functionality and the participation of women and children. The proportion of SBMCs that met our standard for inclusion of women and children remains low, at 14% for women's participation and 4% for children's participation.

School quality: ESSPIN schools had significantly higher quality scores in 2014 than non-ESSPIN schools, and greater levels of intervention were associated with faster improvement in quality between CS1 and CS2. School quality in non-ESSPIN schools stagnated between 2012 and 2014, and there was consequently no significant change in school quality across the state as a whole.

Pupil learning: Schools which have benefited from the ESSPIN interventions tend to have significantly better pupil test results in 2014 than control schools. The state average test scores in Jigawa range from 17% in grade 4 literacy to 29% in grade 2 numeracy, and did not see an improvement on average between CS1 and CS2. Test scores appear to have declined slightly in both ESSPIN and non-ESSPIN schools, although the change is not significant.

In every aspect of school functioning, teaching, and learning outcome measured, ESSPIN schools in Jigawa are performing better than non-ESSPIN schools. In terms of school development planning, SBMC functionality, participation of women and children in SBMCs, and overall school quality, schools that had more ESSPIN intervention also improved faster (or improved while other schools worsened) between 2012 and 2014. Despite these positive findings for ESSPIN, most indicators have not changed significantly over time when we look at Jigawa's schools as a whole. School development planning and some aspects of teacher behaviour improved between 2012 and 2014, while school inclusiveness worsened significantly. Children's test results remain low in both ESSPIN and non-ESSPIN schools.

These two findings – the better quality of ESSPIN schools compared to control schools, and the stagnation over time in the state as a whole – can be reconciled by noting that ESSPIN interventions have been rolled out to around one-third of schools in the state, so the remaining two-thirds have not yet benefited. These results also come in a context of a state that is experiencing threats of violence and conflict. The impact of this tension and insecurity cannot be quantified, but it is assumed to have made it more difficult for ESSPIN intervention to achieve the expected results. The context of rapidly increasing enrolment in the state is also likely to have limited schools' and teachers' ability to deliver better learning outcomes.

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

Positive results in this report include:

- Schools which have received targeted support from ESSPIN tend to be of higher quality than control schools. In addition, ESSPIN schools increased their level of quality between 2012 and 2014, while control schools' quality remained stagnant.
- Pupils in ESSPIN schools score significantly higher results in numeracy and literacy than pupils in control schools.
- Significantly more teachers who have received training under ESSPIN meet the competence standard than teachers in control schools, and some aspects of teacher behaviour in the classroom have improved between 2012 and 2014.
- There was a significant improvement in school development planning in Jigawa between 2012 and 2014.
- Overall school quality is better in ESSPIN than control schools, and has also improved between 2012 and 2014 in ESSPIN schools.
- ESSPIN schools have significantly more functional and inclusive SBMCs than control schools.

Some challenges identified in the report include:

- The proportion of Jigawa's schools meeting the inclusiveness standard fell between 2012 and 2014.
- Teachers' knowledge of benchmarks in the English or mathematics curriculum fell from 57% in 2012 to 21% in 2014, and overall teacher competence has remained stagnant.
- Overall school quality has stagnated in schools that have not yet received ESSPIN intervention.
- Average test results in numeracy and literacy for pupils in grades 2 and 4 appeared to fall between 2012 and 2014 across the state.

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
PTR	Pupil–teacher ratio
SBMC	School-based management committee
SDP	School development plan
SIP	School Improvement Programme

1 Introduction

The aims of the ESSPIN 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 Jigawa 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 Jigawa 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 the 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 by 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 school development plans (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 have received the ESSPIN delivery model.

1.2 ESSPIN in Jigawa State

ESSPIN has been working in partnership with the government of Jigawa State to realise sustainable school improvement since 2009. In Jigawa the ESSPIN intervention began with Phase 1 (pilot) schools – 198 schools, or 9% of all the government primary and junior secondary schools – in 2009/10. The pilot intervention focused on building the skills of teachers, headteachers and SBMCs drawn from members of the local communities. Later the programme was rolled out to additional schools in 2012/13 and 2013/14.

The key school-level interventions in Jigawa's ESSPIN schools (schools which have benefited from the ESSPIN model of support) took place between 2009/10 and 2010/11 and included (ESSPIN 2013b, 2014a):

- Primary school headteachers received training on:
  - academic leadership;
  - school planning;
  - management of teachers; and
  - working with the community.
- Teachers (an average of four teachers in every primary school worked with) received training on:
  - basic literacy teaching (initial reading skills);
  - basic numeracy teaching (number concepts, and addition and subtraction); and
  - use of teaching aids, classroom organisation, and praise.
- The 198 pilot 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. The 303 Phase 2 schools were given a sum of between N20,000 and N50,000 in 2013, based on their size.
- Community members (18 people from each of the ESSPIN schools) 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;
  - child protection and participation;
  - inclusive education and gender; 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 Jigawa 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

The ESSPIN programme has been gradually rolled out to more government primary schools in Jigawa, such that by 2013, 34% of schools, and by the time of the 2014 Composite Survey, 48% of schools had benefited from the full package for at least one year (Table 1; Annex B). The scale of the roll-out of Output stream 3 consisted of the following stages:

- Phase 1 (the pilot phase) consisted of 198 schools. In selecting these schools, the State Universal Basic Education Board selected nine local government education areas (LGEAs) (one from each of the training zones) and then selected approximately 25% of the schools from each of these LGEAs (ESSPIN 2013a). These schools received two consecutive years of support from ESSPIN, and then a year without support; the teacher training interventions resumed again in 2012/13.
- Phase 2 expanded the programme to 303 schools in 2012/13.
- In 2013/14 a further 501 schools joined the ESSPIN programme.

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
Jigawa	10	10	0	15	39	48

Source: Author's 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.

A summary of the characteristics of Jigawa's schools according to the level of ESSPIN intervention is given in Annex A. Generally the schools with more years of ESSPIN intervention are more likely to be urban, have higher enrolment and more teachers, but a lower pupil–teacher ratio (PTR), and are in better condition in terms of seating, blackboards and general infrastructure. This has to be kept in mind in interpreting the following results, as differences in outcomes between ESSPIN and non-ESSPIN schools are likely to reflect these different characteristics as well as changes due to the intervention itself.

According to schools captured in the census, enrolment increased by more than 36%, rising from 427,000 to 584,000 between 2009 and 2013 (Table 2). Enrolment in the schools which were captured in both census increased by over 18%. There were large increases in the PTR and pupil–classroom ratio between 2009 and 2013 (Annex A). In 2013, some 41% of schools in Jigawa had over 50 pupils for each teacher, compared to 29% in 2009. This has to be kept in mind in

interpreting the results in the following sections, as the massive increases in PTR are likely to have made it difficult for schools and teachers to deliver good learning outcomes.

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

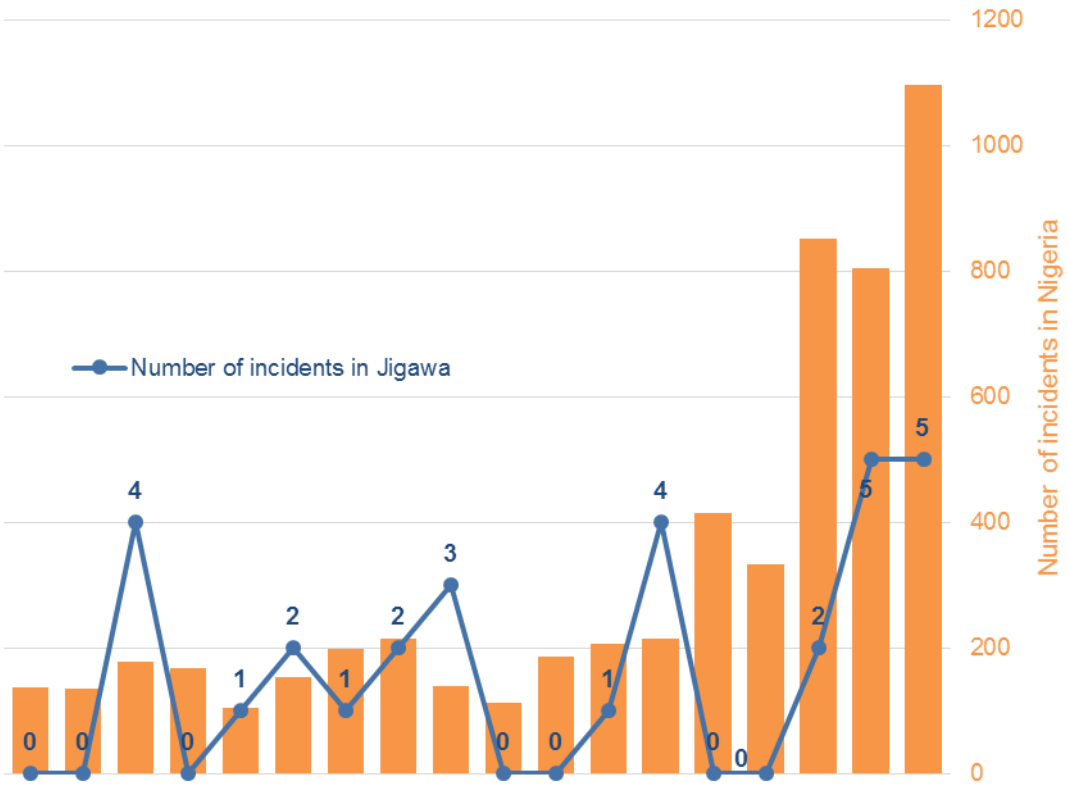
	2009		2013		Enrolment change (%)	Enrolment change (schools found in both censuses only, %)
State	Schools	Enrolment	Schools	Enrolment		
Jigawa	1789	427,180	2157	584,037	36.7	18.5

Note: Enrolment is for primary grades 1–6.

1.4 Conflict in Jigawa

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, and in Jigawa has fluctuated between no incidents and five incidents recorded per year (Figure 1). These five recorded violent events (2014) caused 17 fatalities (Table 3).

Figure 1. Incidents of political violence in Nigeria and Jigawa



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. Jigawa: Political violence: Incidents and fatalities, 2010–2014

Variable	2010	2011	2012	2013	2014
Events	0	0	2	5	5
Fatalities	0	0	1	9	17

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 Jigawa's schools and communities close to the border with Yobe State; members of these communities have heard and witnessed violence and report a resulting influx of people into Jigawa State. In addition, conflict in neighbouring states made access to Jigawa difficult, affecting the support that could be provided. The potential impact of this context should be considered in the reading of this report.

2 Methodology and analysis

2.1 Evaluation strategy

2.1.1 Classifying the amount of ESSPIN intervention

The original evaluation design for ESSPIN relied on maintaining a control group of schools with no intervention, which could be compared to those with a longer history of intervention (Phase 1: roll-out prior to the 2012/13 school year) and to those where intervention had started more recently (Phase 2: roll-out in 2012/13 or 2013/14). For Jigawa we have grouped schools according to the amount of intervention they have had (see Annex B for full details). We focus on schools which had a 'full package' of Output stream 3 interventions in a particular year – leadership training for headteachers, teacher training, and school visits – and treat schools that had less than this full package as control schools. We also assume that there is a one-year lag between ESSPIN intervention and measurable impact, so schools which benefited only in 2013/14 are considered to be part of the control group.

In the following sections we compare schools which had ESSPIN intervention before 2013/14 (labelled 'ESSPIN' or 'intervention schools') to those that either had intervention only in 2013/14, or no intervention at all (labelled 'control schools'). The expectation is that the schools that had intervention prior to 2013/14 should be performing better than those that did not.

However, it is also useful to consider whether schools which had intervention during the relevant period improved faster during 2012 to 2014 than those which did not. This means comparing the schools that had intervention in 2012/13 to those that did not.

To examine 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 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 ESSPIN states 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 Jigawa, 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 schools that had intervention prior to 2013/14 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 schools that had intervention in 2012/13 and over CS1 and CS2. 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.

2.2 Sampling, coverage and weights

In Jigawa the sample allocation for CS1 was 105 schools, giving 35 each in control, Phase 1 and Phase 2 categories of schools. In total 103 were able to be sampled, and these same schools were visited for CS2, along with two additional schools to bring the sample size back up to 105 (Megill, 2014) – during fieldwork two of the intended schools were not visited, one because of security concerns and another because it was found to be ineligible for the survey.[1] 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

Intervention	CS1 sample	CS2 sample	Population
No intervention/intervention in only 2013/14 (minimum)	32	32	1569
Since 2012/13 (medium)	36	36	303
In pilot – 2009/10 and 2010/11 (medium)	35	35	198
Total	103	103	2070

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

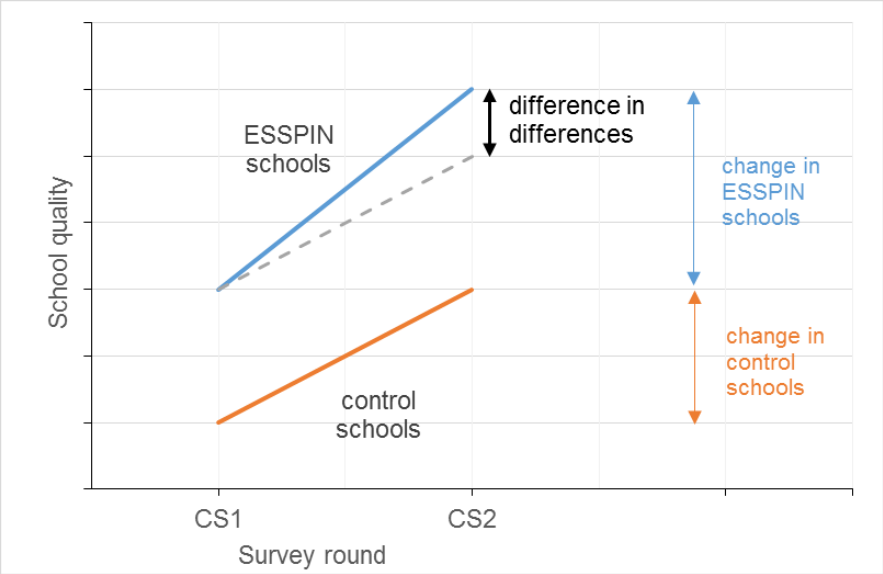
In each school the headteacher was interviewed, as was the SBMC chairperson.

[1] This school was a special school for deaf children with an adapted curriculum. It was therefore not comparable with other schools in Jigawa.

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 – in Jigawa, this means the group where intervention begun in 2012/13 – 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 the 2012/13 cohort of ESSPIN schools. Is the change over time greater in the 2012/13 ESSPIN schools than in the other 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.

Teachers within each school were sampled from the population present in the school on the day of the survey visit and who taught grades 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 94% of the intended sample of pupils was included, and 66% of teachers. The actual numbers of schools, teachers and students sampled is given in Table 5.

Table 5. Jigawa: Sample coverage in CS2

	Schools		Teachers			Pupil tests			
	Intended sample	Actual	Interview	Less. Obs.	Tests	L2	L4	N2	N4
Jigawa	105	103	430	425	415	393	399	396	398

Note: Throughout this report, *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 grade 2 and grade 4 were given tests in either literacy or numeracy.

Two indicators of aggregate learning outcomes are used in this note. 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 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 and CS2. 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 benefited from ESSPIN or did not. Here we expect the results to be better for schools which benefited from ESSPIN. If a school entered the programme in 2013/14, we count them 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 2012/13, with those which did not, to see whether they improved more or faster between 2012 and 2014 than the other schools. Here we expect the supported schools to improve relative to other schools over the two years.

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 Jigawa's teachers in CS1 and CS2. (The fifth criterion is not available here as teacher tests were not conducted as part of CS1.) Use of teaching aids and use

of praise during classes have both improved significantly. Teachers' knowledge of the English or maths curriculum has significantly dropped in Jigawa, from 57% to 21%. Overall in Jigawa, the proportion of teachers meeting the (CS1) competence standard is now 70%, an increase from 62% in CS1, though this change is not statistically significant. 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 also not significantly different in CS2 than in CS1.

Table 6. Jigawa: Teacher competence in CS1 and CS2

	CS1	CS2	
(1) Knowledge of Eng/maths curriculum	57.1	21.3	-
(2) Use of 1+ teaching aid	89.4	99	+
(3) Praise more than reprimand	63.2	83.6	+
(4) Assigns 2+ ind./group task	47	58.6	
Competence score (CS1 version)	65.2	71.4	
Teacher competence standard (CS1)	62.2	69.8	

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 evidence suggests that teachers in ESSPIN schools may be more competent than teachers in control 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. Across the individual criteria, and the literacy and numeracy tests, there are no significant differences between teachers in ESSPIN schools (including those who individually received training) and those not in ESSPIN schools. However, performance was higher for ESSPIN-trained teachers in the use of teaching aids other than the blackboard, in the use of praise, and on the scores in the literacy and numeracy tests. The results find that 19% of teachers who benefited from ESSPIN training meet the new competency standard, which is significantly more than the 3% of teachers who meet this standard in non-ESSPIN schools.

The literacy and numeracy scores for teachers in Jigawa are low. Teachers who were trained under ESSPIN received the highest average scores; these were 36% for literacy and 57% for numeracy. Less than one-quarter of teachers passed the literacy and numeracy test. Teachers' difficulties with the content, particularly the use of English, will undoubtedly have an impact on pupil learning, especially the pupil test results for literacy (see section 3.7).

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

	(1) Non-ESSPIN	(2) ESSPIN school		(3) ESSPIN-trained	
Knowledge of Eng/maths curriculum	21.3	23.7		20	
Use of 1+ teaching aid	99.5	97.7		98.2	
Use of 1+ teaching aid excl read/write/copy from blackboard	76.1	70		90	
Praise more than reprimand	81	87.1		90.3	
Assigns 2+ ind/group task	59.5	53.4		60.1	
Literacy score (%)	31.6	34.2		36.3	
Numeracy score (%)	51.5	54		56.7	
Passes literacy and numeracy test	10	13.7		23.7	
Competence score (CS1 version)	70.6	73.8		72.4	
Teacher competence standard (CS1)	66.2	79.3		74.6	
Competence score (CS2 version)	57.9	61.6		64.8	
Teacher competence standard (CS2)	2.6	8.3		19	+

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 ($p < .05$).

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. However, there is some suggestion that while teachers who are not in ESSPIN schools (column 1) improved their competency on average, the average scores of teachers who were individually trained by ESSPIN (column 3) actually fell. An alternative method of difference in differences analysis, using regression with continuous variables, is shown in Annex D. In the regression we also adjusted for how long a teacher had been at the school, to make sure we were including only those teachers who had been at the school long enough to have benefited from the training. With these results we also found no significant difference from teachers in non-ESSPIN schools.

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

Teacher competence scores (CS1 version)	(1) Non-ESSPIN	(2) ESSPIN school	(3) ESSPIN-trained
CS1	64.2	59.4	78
CS2	70.6	73.8	72.4
Difference	6.4	14.4	-5.7

Note: * indicates a significant difference in differences compared to the non-ESSPIN schools ($p < .05$).

Overall in Jigawa the findings suggest that average teacher competence did not change significantly between 2012 and 2014. In 2014, teachers who have received training through ESSPIN at some point were more likely to meet the competence standard than those who had not been trained. However, there is no evidence to suggest that teachers in Jigawa improved their competence faster between 2012 and 2014 if they had received individual training through ESSPIN in recent years.

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 Jigawa, teachers were almost twice as strong in reading as they were in writing, and grammar was also a stronger area (Figure 2). In mathematics, teachers performed better in 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 Jigawa than the English questions.

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

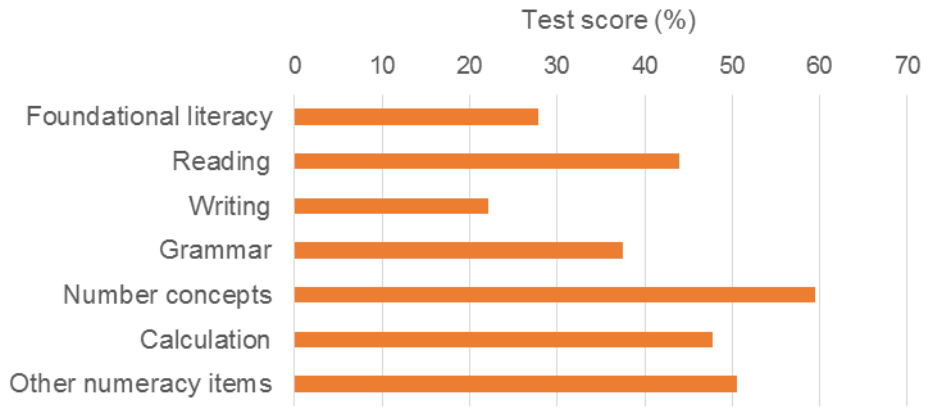
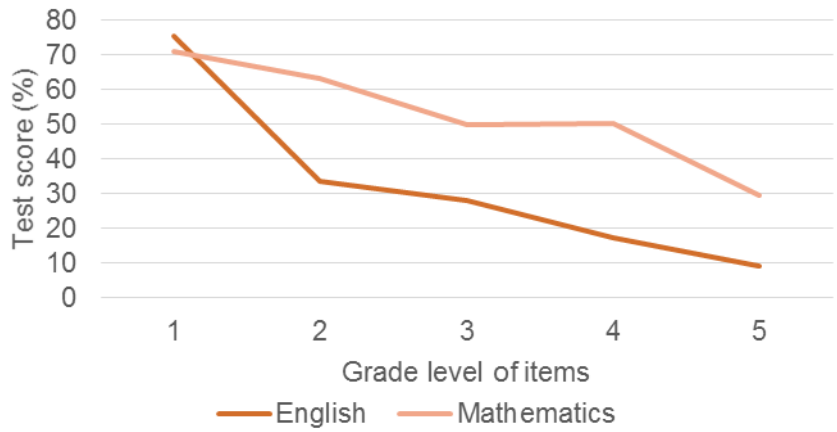


Figure 3. Jigawa: Teacher test scores by grade



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 did not change significantly between CS1 and CS2 (Table 9): 9% of headteachers met the overall standard in CS2, which was a decline from 17% in CS1 (although the decline was not statistically significant). Similarly the average number of criteria met fell marginally, from 3.3 to 2.9 (out of 7). Performance on the individual criteria was varied; headteachers holding lesson observations, professional development meetings and schools with a clear opening time all increased, but the change was not significant. The only significant change was a fall in the proportion of headteachers taking action on teacher attendance.

Table 9. Jigawa: Headteacher effectiveness in CS1 and CS2

	CS1	CS2	
(1) Lesson observations	6.8	10.1	
(2) Professional development meetings	6.7	20.3	
(3) Action on teacher attendance	80	39.9	-
(4) Clear opening time	41.5	51.2	
(5) In class on time	52.9	32.1	
(6) Appropriate morning break	89.8	73.4	
(7) Appropriate lesson length	61.5	60.4	
Number of criteria fulfilled (/7)	3.3	2.9	
Effective headteacher (5/7 criteria met)	16.9	9.1	

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

Focusing on the 2014 data, headteachers who received leadership training from ESSPIN are more effective than those who have not received training (Table 10), with 26% meeting the effectiveness criteria, as opposed to 4% in non-ESSPIN schools. The proportion of headteachers conducting lesson observations was higher for the ESSPIN group. Surprisingly, however, lessons were less likely to be the appropriate length in schools where the headteachers had received ESSPIN leadership training.

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

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) Lesson observations	2.6	35.1	+
(2) Professional development meetings	17.2	30.9	
(3) Action on teacher attendance	37	49.4	
(4) Clear opening time	47.2	64.2	
(5) In class on time	25.4	54.4	
(6) Appropriate morning break	76.1	64.7	
(7) Appropriate lesson length	66.8	38.8	-
Number of criteria fulfilled (/7)	2.8	3.4	
Effective headteacher (5/7 criteria met)	4.1	26	+

Note: + indicates a significant difference from the results in non-ESSPIN schools ($p < .05$).

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 worsened more slowly) than comparators. The results suggest there was no significant difference between the change in the number of criteria met by headteachers in schools benefiting from ESSPIN and non-ESSPIN schools (Table 11). We also adjusted for the year that the headteacher was appointed to his or her current school to account for transfers, and still found no significant difference.[2] Similar results were found when using a regression analysis method, which is shown in Annex D.

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

Number of criteria met (/7)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	3.2	3.4
CS2	2.9	3.1
Difference	-0.4	-0.3

Note: * indicates a significant difference in differences compared to the non-ESSPIN schools ($p < .05$).

In summary, headteacher effectiveness does not appear to have changed significantly overall in Jigawa, although it is better in ESSPIN schools than non-ESSPIN ones. 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). There was a significant improvement in school development planning in Jigawa between CS1 and CS2, particularly in terms of evidence of self-evaluation, having a current SDP available, and carrying out at least four activities from the plan (Table 12). The results suggest school development planning is still low in Jigawa, as the average number of criteria met is 0.7 out of five (up significantly from 0.3), and 6% of schools meet the overall standard (up from 1%).

[2] This is shown explicitly in the regression table in Annex D but similar results are obtained for the comparison of means method when we adjust the intervention categories to take into account the date when the headteacher was appointed to his or her current school.

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. Jigawa: SDP effectiveness in CS1 and CS2

	CS1	CS2	
(1) Written evidence of school self-evaluation process	9.6	23.1	+
(2) SDP available	9.9	21.7	+
(3) SDP contains 3+ activities to strengthen teaching and learning	2.3	6.2	
(4) Evidence that 4+ activities from SDP carried out	1	6.4	+
(5) Cashbook up-to-date	5.9	11.9	
Number of SDP criteria fulfilled (/5)	0.3	0.7	+
School meets effective school development planning standard (%)	1	6.2	+

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

Focusing on CS2, across all of the five criteria, ESSPIN schools are doing significantly better than non-ESSPIN schools where we did not expect an improvement (Table 13). The proportion of schools with evidence of self-evaluation criteria was 64% in ESSPIN schools and only 11% in control schools. There is still room for improvement in the ESSPIN schools, with less than a quarter of schools planning activities to improve teaching and learning, or having evidence of executing their SDP activities.

Around one in four schools (24%) that have received ESSPIN intervention meet the school development planning standard, which is low but significantly better than only 1% of non-ESSPIN schools meeting the standard. ESSPIN schools on average meet two out of five criteria; again, this is still low but it is significantly better than non-ESSPIN schools, which fulfil only 0.3 of the criteria on average.

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

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) Written evidence of school self-evaluation process	10.7	64.2	+
(2) SDP available	10.7	58.1	+
(3) SDP contains 3+ activities to strengthen teaching & learning	1.6	21.5	+
(4) Evidence that 4+ activities from SDP carried out	0.8	25.2	+
(5) Cashbook up-to-date	4.8	35.3	+
Number of SDP criteria fulfilled (/5)	0.3	2	+
School meets effective school development planning standard	0.8	24	+

Note: + indicates a significant difference from the results in non-ESSPIN schools ($p < .05$).

To assess whether ESSPIN schools improved faster than non-ESSPIN schools in terms of school development planning, we compare the change in average criteria met for the two groups (Table 14). The results find a positive significant difference in the change between ESSPIN and non-ESSPIN schools, meaning that ESSPIN schools improved faster than control schools. It is also interesting that in CS1 the ESSPIN schools seem to have been less effective at school development planning, but to have rapidly overtaken other schools, which could be attributed to the more recent school improvement interventions. The regression method of analysis, shown in Annex D, supports these results.

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

Number criteria met (/5)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	0.3	0.1
CS2	0.4	2
Difference	0.1	1.9*

Note: * indicates a significant difference in differences compared to the non-ESSPIN schools ($p < .05$).

In summary, school development planning is improving in Jigawa, and is better in ESSPIN than in non-ESSPIN schools, although in general the performance is low, with only 6% of schools, on average, meeting the standard in 2014. There is evidence to suggest that school development planning in the schools which had more ESSPIN intervention improved faster than in the 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 Jigawa generally there was a decline in the proportion of all schools meeting the inclusiveness criteria between 2012 and 2014, with all indicators falling, except for the indicator of activities to improve access in the SDP (Table 15). These reductions led to the overall proportion of schools meeting the standard falling significantly (from 20% to 5%).

Box 7. Standard for school inclusiveness (meeting the 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. Jigawa: School inclusiveness in CS1 and CS2

	CS1	CS2	
(1) 3+ actions to improve attendance	66.2	30.9	-
(2) 2+ activities in SDP to improve access for disadvantaged children	1.2	13.5	+
(3) >50% of teachers use 2+ assessment methods	59.5	27.8	-
(4) >50% of teachers spatially inclusive and >50% are gender inclusive	31.6	15.8	
Number of inclusiveness criteria fulfilled (/4)	1.6	0.9	-
Inclusiveness score	69.8	57.8	-
School partially met inclusiveness standard (2–4 criteria out of 4)	49.3	27	-
School fully met inclusiveness standard (3–4 criteria out of 4)	19.8	5.1	-

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

Focusing on CS2 schools, more ESSPIN schools met the criteria on having activities to improve access in their SDP (36%) compared with non-ESSPIN schools (7%). The average inclusiveness score was higher in ESSPIN schools than in non-ESSPIN schools; the difference was small but statistically significant at 60% compared with 57%.

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

	(1) Non-ESSPIN	(2) ESSPIN	
Inclusiveness criteria			
(1) 3+ actions to improve attendance	31.9	27.3	
(2) 2+ activities in SDP to improve access for disadvantaged children	6.7	36.4	+
(3) >50% of teachers use 2+ assessment methods	25.8	34.8	
(4) >50% of teachers spatially inclusive and >50% are gender inclusive	16.5	13.5	
Overall inclusiveness standard			
Number of inclusiveness criteria fulfilled (/4)	0.8	1.1	
Inclusiveness score	57.1	60.4	+
School partially met inclusiveness standard (2–4 criteria out of 4)	23.1	39.8	
School fully met inclusiveness standard (3–4 criteria out of 4)	**4.6**	**6.7**	
Detailed			
Number of actions to improve attendance	2.3	2	
Number of activities on access for disadvantaged children	0.2	1.2	+
Average number of assessment methods used	0.4	0.6	
Average number of zones participating in lessons	3.6	3.9	
Average gender equity score (0=completely unequal, 100=perfectly equal)	78.6	86.6	+

Note: + indicates a significant positive difference between non-ESSPIN and ESSPIN schools. 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.

Difference in difference analysis of schools in Jigawa depending on the level of ESSPIN intervention has been carried out by comparing the change in inclusiveness score (Table 17). The results suggest that both groups of schools became less inclusive between 2012 and 2014. Although the results are not statistically significant, there is some suggestion that the schools with more ESSPIN activities worsened less (or more slowly) than other schools over the period. Using a different method of analysis, we found similar results – which are shown in Annex D.

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

Inclusiveness score	(i) Non-ESSPIN	(ii) ESSPIN
CS1	69.9	69.8
CS2	57.4	60
difference	-12.5	-9.8

Note: * indicates a significant difference in differences compared to the 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).

In Jigawa, out of the 103 schools sampled in CS1, 93 had SBMCs (90%) (Table 18). In CS2 all of the schools sampled had SBMCs, showing an improvement in the establishment of SBMCs. This does not mean that all the SBMCs are functional or inclusive, however, or that they have all received ESSPIN Output stream 4 support. Jigawa did not conduct women's and children's participation training in any schools until 2012/13, so we cannot expect a 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. Jigawa: Sample size of schools with SBMCs

	CS1	CS2
Schools sampled in Jigawa	103	103
Schools with SBMCs sampled in Jigawa	93	103

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 Jigawa there was little difference between average levels of functionality in CS1 and CS2, apart from a significant improvement in SBMC networking (criterion

4) and a decline in the number of SBMC chairs visiting the school regularly (criterion 9) (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.

Two additional criteria related to the inclusiveness of SBMCs (not included in the CS1 report) are also examined in this section: 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. There was a significant increase in SBMCs raising issues of exclusion in Jigawa.

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) The SBMC chair visited the school at least three times from the start of the current school year (written evidence).

[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.

Table 19. Jigawa: SBMC functionality in CS1 and CS2

	CS1	CS2	
(1) 2+ meetings this school year	24.7	20.6	
(2) Conducted awareness-raising	51.8	40.8	
(3) Addressed exclusion	31.9	20.3	
(4) Networked with CBOs/institutions/other SBMCs	27.1	61.9	+
(5) Interacted with LGEA	22.3	16.4	
(6) Has women's committee	20.6	15.3	
(7) Has children's committee	15	14	
(8) Contributed resources for school	39.4	41.7	
(9) Chair visited school 3+ times	41.6	13.2	-
Schools meeting functioning SBMC standard	20.4	22.3	
Number of SBMC functionality criteria met (/9)	2.8	2.6	
Additional criteria			
Action for commonly excluded groups	22.8	21.2	
Raised issue of children's exclusion	3.9	20.3	+

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 Jigawa's ESSPIN schools performed more effectively on all the SBMC functionality criteria, and significantly so on seven out of nine of them. Notably, the proportion of schools meeting the overall standard was far higher in ESSPIN schools – at 65% – compared with non-ESSPIN schools (8%). Schools with ESSPIN interventions also had better results than the control schools on the promoting inclusiveness criteria.

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

	(1) Non-ESSPIN	(2) ESSPIN	
(1) 2+ meetings this school year	8.1	59.6	+
(2) Conducted awareness-raising	34.2	61.4	+
(3) Addressed exclusion	14.7	38.1	+
(4) Networked	56.7	78.3	
(5) Interacted with LGEA	12.7	26.9	
(6) Has women's committee	2.2	56.2	+
(7) Has children's committee	0.8	55.2	+
(8) Contributed resources for school	28.9	80.6	+
(9) Chair visited school 3+ times	3.1	43.9	+
Standard G: functioning SBMC	8.3	64.8	+
Number of SBMC functionality criteria met (/9)	1.7	5	+
Additional criteria			
Action for commonly excluded groups	16.8	35	
Raised issue of children's exclusion	14.7	38.1	+

Note: + indicates a significant difference from the results in non-ESSPIN schools (p < .05).

We have seen that the overall level of SBMC functionality did not significantly improve between the surveys, and we are interested to know whether SBMCs which received more ESSPIN support improved faster than other schools. The comparison of means analysis (Table 21) finds that SBMCs in ESSPIN schools improved significantly more than non-ESSPIN schools (which actually worsened) between CS1 and CS2. The regression method supports this finding and is shown in Annex D.

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

Number criteria met (/9)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	3	2.4
CS2	2	4.9
Difference	-1	2.6*

Note: * indicates a significant difference in differences compared to the 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 for the improving of the participation of women and children, which included establishing sub-committees, was rolled out after initial SBMC training and mentoring. In Jigawa this was not until 2012/13, and in only a subset of schools (see Annex C), so we might not expect to see much impact on the state averages for women and children's participation at the time of the CS2.

In fact in Jigawa there was a significant improvement between CS1 and CS2 in the number of SBMC women's committees which met (Table 22). The number of schools meeting the overall standard for SBMC women's inclusiveness increased from 8% to 14%, though this was not found to be statistically significant.

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]

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

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

	CS1	CS2	
(1) At least one woman attended 2+ meetings (%)	17.6	9.5	
(2) Female member raised an issue (%)	22.3	24.3	
(3) Issue raised by female member led to action (%)	19.3	14.1	
(4) Women's committee met (%)	5.9	27.4	+
Number of criteria met	0.6	0.8	
Meets standard (3/4 criteria)	7.9	13.6	

Note: 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).

Women's inclusiveness was found to be significantly higher in ESSPIN schools than non-intervention schools for all the criteria in CS2 (Table 23). Overall 43% of ESSPIN schools in Jigawa met the standard but none of the non-ESSPIN schools did. No non-ESSPIN schools met the standard because none of them met the criteria on women attending meetings, and therefore also could not have an issue raised by female members that led to action.

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

	(1) Non-ESSPIN	(2) ESSPIN	
(1) At least one woman attended 2+ meetings (%)	0	31.7	+
(2) Female member raised an issue (%)	10.8	53	+
(3) Issue raised by female member led to action (%)	0	41.9	+
(4) Women's committee met (%)	10.4	59.4	+
Number of criteria met	0.1	1.9	+
Meets standard (3/4 criteria)	0	42.9	+

Note: + indicates a significant difference from the results in non-ESSPIN schools (p < .05).

Difference in differences analysis finds a significant treatment effect, as would be expected, implying that schools receiving ESSPIN intervention in 2012/13 did in fact improve their SBMC women's inclusiveness faster over the period than other schools (Table 24).

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

Number criteria met (/4)	(1) Non-ESSPIN	(2) ESSPIN
CS1	0.6	0.7
CS2	0.3	1.8
Difference	-0.2	1.1*

Note: Schools that did not have SBMCs at all in CS1 are excluded from the analysis; * indicates a significant difference in differences compared to the 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 in the standard on SBMC children's inclusiveness.

As mentioned above, the intervention support for children's participation was rolled out to only a selection of schools in 2012/13, so we cannot expect a large impact on the indicators. It may be unsurprising therefore that children's participation in SBMCs is low in Jigawa and that there was little change between CS1 and CS2. The only significant change was a reduction in the number of SBMCs in which a child attended at least two meetings. The proportion of schools' SBMCs meeting the overall standard is very low at only 4.1%.

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]

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

	CS1	CS2	
(1) Child attended 2+ meetings (%)	12.5	4.3	-
(2) Child raised an issue (%)	16	19.4	
(3) Issue raised by child led to action (%)	12.3	7.9	
(4) Children's committee met and has a trained facilitator (%)	1.9	4.5	
Number of criteria met	0.4	0.4	
Meets standard (3/4 criteria) (%)	3.5	4.1	

Note: 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, there are positive differences between ESSPIN and non-ESSPIN schools on all the indicators (Table 26), significantly so in terms of children attending SBMC meetings, raising an issue in the SBMC, and the issue raised leading to action. Overall, 17% of ESSPIN schools met the standard for SBMC children's inclusiveness while none of the non-ESSPIN schools did.

Table 26. Jigawa: SBMC children's inclusiveness in CS2, between intervention groups

	(i) Non-ESSPIN	(ii) ESSPIN	
(1) Child attended 2+ meetings (%)	0	17.6	+
(2) Child raised an issue (%)	10.4	46.8	+
(3) Issue raised by child led to action (%)	2.3	26.7	+
(4) Children's committee met and has a trained facilitator (%)	2.4	11.1	
Number of criteria met	0.2	1	+
Meets standard (3/4 criteria) (%)	0	16.9	+

Note: + indicates a significant difference from the results in non-ESSPIN schools ($p < .05$).

[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.

The difference in differences analysis finds evidence that ESSPIN schools did improve faster between 2012 and 2014 than non-ESSPIN schools (which may in fact have worsened) (Table 27).

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

Number of criteria fulfilled (/4)	(i) Non-ESSPIN	(ii) ESSPIN
CS1	0.4	0.4
CS2	0.2	0.9
Difference	-0.1	0.5*

Note: Schools that did not have SBMCs at all in CS1 are excluded from the analysis; * indicates a significant difference in differences compared to the non-ESSPIN schools ($p < .05$).

Overall, the results for Jigawa show that schools benefiting from ESSPIN had significantly more inclusive SBMCs and improved their participation from women and children more over the last two years than control schools. However, this was not enough to raise the overall average for the state, which reflects a challenging situation with regard to the inclusion of women and children in SBMC activities, and to their declining participation in the control schools.

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). The survey results suggest there has been a positive but not significant change in the proportion of Jigawa's schools that meet the overall school quality standard between 2012 and 2014. Less than 5% of schools fulfilled at least three out of four individual standards in CS2 (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 (though not significant) increase in this quality score (from 36% in CS1 to 40% in CS2).

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. Jigawa: School quality in CS1 and CS2

	CS1	CS2	
Meets three or four standards (CS1 version)	2.2	4.8	
Quality score (CS1 version)	36	39.7	

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

Within CS2, around a fifth of ESSPIN schools meet the overall school quality standard (CS1 version), whereas almost no other schools meet the standard (Table 29). This shows a significantly higher level of quality in the ESSPIN schools than in the control schools. When we look at the stricter definition of school quality, which includes teachers' results in literacy and numeracy tests, only 5% of the ESSPIN schools meet the standard. This indicates that the poor grasp of literacy in English is dragging down the overall level of school quality.

Using our continuous indicator of school quality, scores are significantly higher in ESSPIN than non-ESSPIN schools, at around 55% as opposed to 31%.

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

	(1) Non-ESSPIN	(2) ESSPIN	
Meets three or four standards (CS1 version)	0.2	20.3	+
Meets three or four standards (CS2 version)	0	5	
Quality score (CS1 version)	34.6	53.9	+
Quality score (CS2 version)	30.9	55.8	+

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 ($p < .05$).

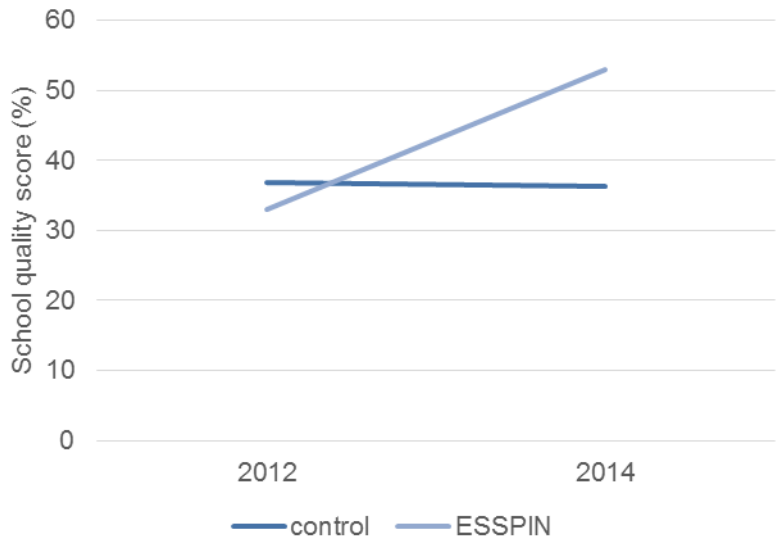
Difference in differences analysis of the change in mean quality score finds that ESSPIN schools improved significantly more or faster than control schools (Table 30). This suggests that the impact of ESSPIN activities was positive on the beneficiary schools between 2012 and 2014. This impact can be seen clearly in Figure 4, where the non-ESSPIN schools have remained at the same level of quality between 2012 and 2014, but ESSPIN schools have overtaken them steeply. Using a different method of analysis we find similar results, as shown in Annex D.

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

School quality score	(1) Non-ESSPIN	(2) ESSPIN
CS1	36.9	33
CS2	36.3	53
Difference	-0.6	20.1*

Note: * indicates a significant difference in differences compared to the non-ESSPIN schools ($p < .05$).

Figure 4. Jigawa: 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 that received intervention in 2012/13 compared with those that did not. 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 Jigawa range from 17% in grade 4 literacy to 29% in grade 2 numeracy (Table 31). Results are generally higher in grade 2 than in grade 4, and higher in numeracy than in literacy. Pupil test scores saw a negative change in all four subjects in Jigawa between 2012 and 2014, with a significant decline in grade 4 numeracy. Very few pupils met the logframe indicator but these are designed with very specific criteria.

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

	Test	CS1	CS2	
Test score (%)	L2	24.4	19	
	L4	20.9	17.3	
	N2	33.4	28.9	
	N4	27.6	21.4	-
Logframe indicator (%)	L2	0.3	0	
	L4	1.8	0.1	
	N2	4.4	1.2	
	N4	1.3	0	

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

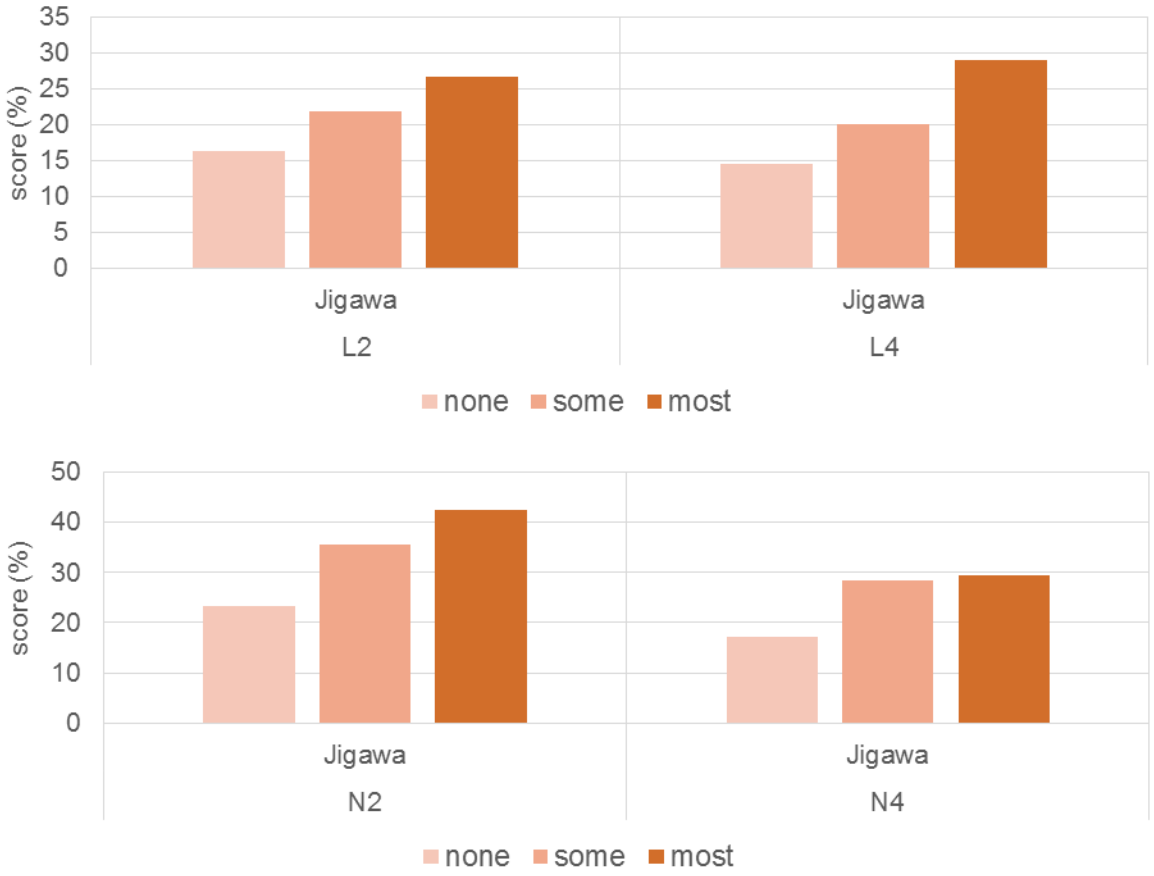
In 2014, pupils who had been taught in ESSPIN schools had significantly higher test scores than those in control schools in all the assessments (Table 32). The difference in test scores according to how much ESSPIN intervention a school had received can be clearly seen in Figure 5, with average scores increasing with the level of ESSPIN intervention. In terms of meeting the logframe indicators, no difference was found between the ESSPIN and non-ESSPIN schools.

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

	Test	(i) Non-ESSPIN	(ii) ESSPIN	
Test score (%)	L2	16.4	23.2	+
	L4	14.5	22.5	+
	N2	23.3	37.7	+
	N4	17.3	28.7	+
Logframe indicator (%)	L2	0	0	
	L4	0	0.3	
	N2	1.7	0.4	
	N4	0	0	

Note: + indicates a significant difference from the results in non-ESSPIN schools ($p < .05$).

Figure 5. Jigawa: Test scores by ESSPIN intervention group



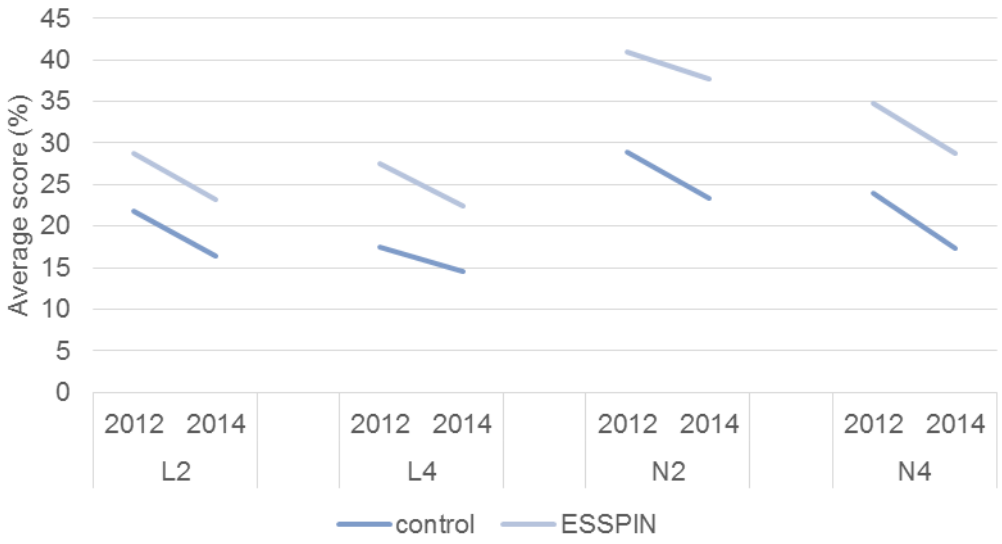
Turning to the question of whether the ESSPIN activities had an impact on change in pupil test scores between 2012 and 2014, we find there was no evidence of a significant impact from the intervention. Generally, average pupil test scores fell between CS1 and CS2 in all types of schools, as seen by the negative differences in the comparison of means method (Table 33). This is shown visually in Figure 6, where test scores fell for both non-ESSPIN and ESSPIN schools across the four subjects. There is evidence that schools benefiting from ESSPIN had higher test scores anyway, particularly in grade 4 literacy and numeracy, confirming the results in Table 32. However, the difference in the change in means is not significant, so it does not appear that ESSPIN lessened the fall in test results in Jigawa's schools between 2012 and 2014. A possible explanation is that there was a lag in impact. This would mean that the schools that received ESSPIN intervention, especially during 2012-13, have not yet felt the effect of the activities on pupil learning.

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

Pupil test score		(i) Non-ESSPIN	(ii) ESSPIN
L2	CS1	21.8	28.7
	CS2	16.4	23.2
	Difference	-5.4	-5.4
L4	CS1	17.5	27.6
	CS2	14.5	22.5
	Difference	-3	-5.1
N2	CS1	28.9	40.9
	CS2	23.3	37.7
	Difference	-5.6	-3.2
N4	CS1	24	34.8
	CS2	17.3	28.7
	Difference	-6.7	-6.1

Note: * indicates a significant difference in differences compared to the non-ESSPIN schools ($p < .05$).

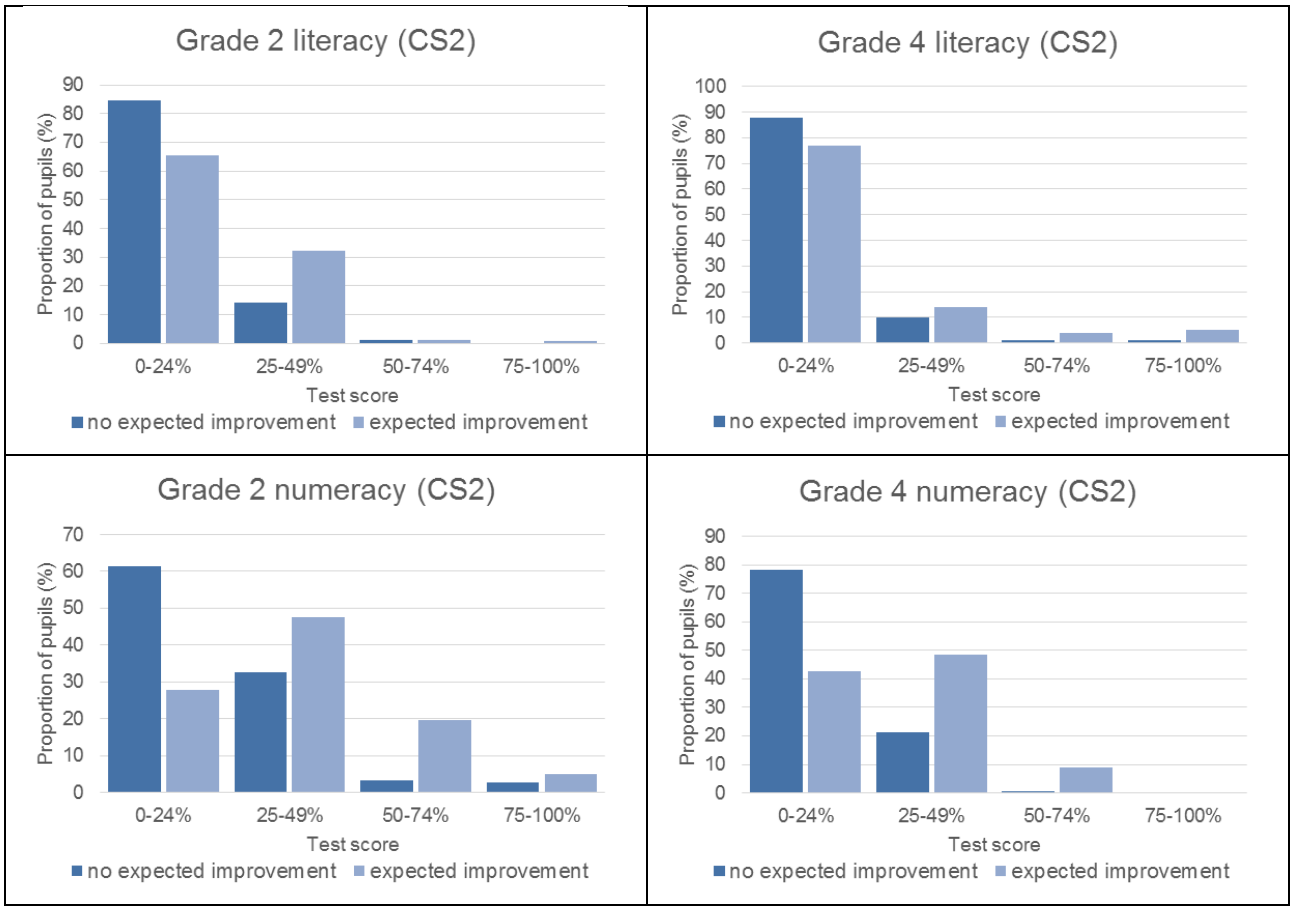
Figure 6. Jigawa: Pupil test scores in ESSPIN and control schools, in 2012 and 2014



3.7.2 Distribution of test scores and sub-scale scores

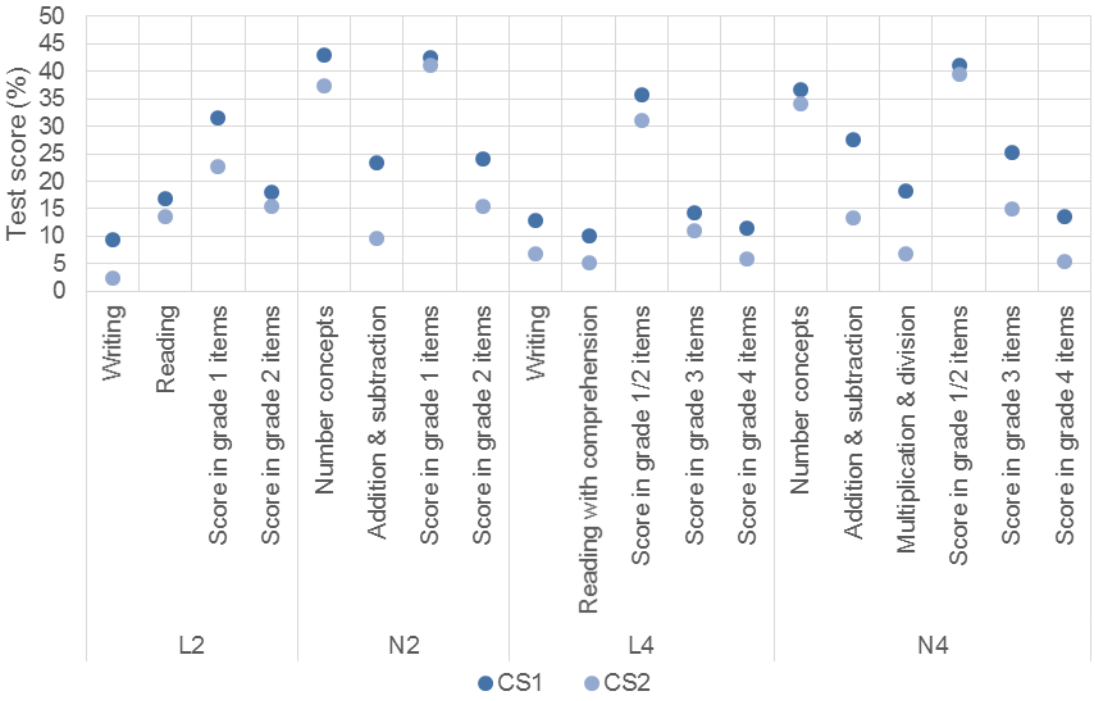
In the 2014 survey, in all four subjects, a very low proportion of pupils scored over 50% (Figure 7). Looking at the distribution across four bands and for schools according to the expected impact from ESSPIN, the results suggest that pupils in ESSPIN schools (labelled 'expected improvement') do have better test results than those in non-ESSPIN schools (labelled 'no expected improvement'). Over 60% of pupils in non-ESSPIN schools scored results below 25%, but a lower proportion of pupils from ESSPIN schools had results in this band. In the numeracy results particularly we see the results we would want: there is a shift to the right, with more of the 'ESSPIN schools' pupils scoring in the higher bands. Even so, there are very few pupils scoring in the 75–100% range.

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



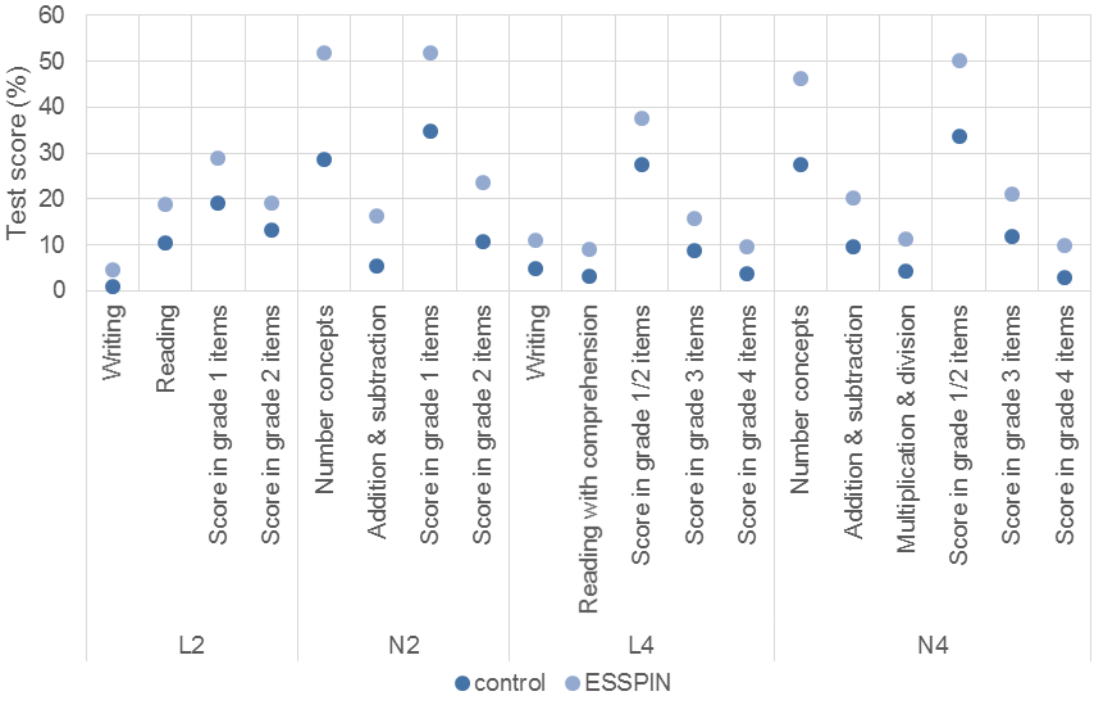
Average pupil test scores have fallen between CS1 and CS2 in all sub-scale areas of the tests (Figure 8). For example, in numeracy in both grades 2 and 4, there was a fall in results in addition and subtraction, and multiplication and division.

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



Across all 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), and the difference was statistically significant. Generally, this difference was quite notable, for example with around a 20 percentage point difference in number concepts and grade 1 items in numeracy (in both grades 2 and 4), and a five to 10 percentage point difference in other areas.

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



4 Conclusion and implications of survey findings for Jigawa'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 Jigawa are:

- In every aspect of school functioning, teaching, and learning outcome measured, ESSPIN schools in Jigawa are performing better than non-ESSPIN schools. This is consistent with ESSPIN having an impact on school quality, although the difference could also be partly explained by the fact that the schools selected for ESSPIN are disproportionately urban and have better infrastructure.

- In terms of school development planning, SBMC functionality, participation of women and children in SBMCs, and overall school quality, schools that had more ESSPIN intervention also improved faster (or improved while other schools worsened) between 2012 and 2014. This provides stronger evidence that better outcomes in ESSPIN schools can be attributed to the intervention itself rather than to pre-existing differences in school characteristics.

- Despite these positive findings for ESSPIN, most indicators have not changed significantly over time when we look at Jigawa's schools as a whole. School development planning and some aspects of teacher behaviour improved between 2012 and 2014, while school inclusiveness significantly worsened. These findings reflect the limited roll-out of ESSPIN, which had reached around one-third of schools in the state prior to 2013/14. The additional roll-out during 2013/14 is unlikely to have had an effect by the time of our survey.

- Children's test results remain low in both ESSPIN and non-ESSPIN schools, and may be worsening. This suggests that better school functioning in schools that received ESSPIN intervention has yet to translate into better pupil learning outcomes.

These findings suggest that ESSPIN is helping to bring up (or at least lessen any fall in) the quality of schools in Jigawa. However, non-ESSPIN schools continue to stagnate or worsen. Potential explanations for this may include very rapid rises in enrolments in Jigawa's schools: in 2013, over 40% of schools in the state had more than 50 pupils for each teacher. Increased pupil numbers would also help to explain why better school quality has not translated into better learning outcomes: it would be difficult for teachers to deliver good lessons, or use newly acquired techniques, when faced with such large and expanding classes.

Teachers' subject knowledge is likely to be an additional 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 only answer around 30% of mathematics items and 10% of English items correctly. 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.

An additional explanation for the stagnation in the quality of non-ESSPIN schools, and in learning outcomes generally, is the level of conflict and violence experienced in the state, as highlighted in section 1.2. The conflict in neighbouring Yobe has caused an influx into Jigawa of displaced people seeking safety, who put a strain on public services, including education. It is likely that the newly enrolled children have come from a fractured experience of education and so have relatively low

learning outcomes, dampening any potential impacts of ESSPIN. There has also been some level of violence within Jigawa State itself.

These threats and pressures on the communities in Jigawa are likely to have an ongoing negative impact on education and many of the indicators measured here. SBMCs might be less likely to meet and function properly, while the stress on teachers and heads might mean they are less effective. These conditions – along with high absenteeism for teachers and pupils – will have had an effect on learning outcomes. Many of the mechanisms though which a school improvement programme can lead to improved learning outcomes may not work in a situation of violence and insecurity.

Despite the inconclusive overall trend for Jigawa, the result that schools where ESSPIN worked did improve more (or worsen less) than other schools is very promising for Jigawa's ESSPIN programme. It provides some support for ESSPIN's theory of change and the associated results chain, in particular that the SIPs do lead to an improvement in school quality. While we do not see a significant impact on learning outcomes, it could be the case that the time lag of the intervention is longer than expected and so learning gains have not yet been realised. The next Composite Survey in 2016 will help test this and see whether learning outcomes eventually do improve.

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) Jigawa State report of ESSPIN composite survey 1 (2012). Report number JG201

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 (2014a) SSO Synthesis Report 2014

ESSPIN (2014b) 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 Jigawa's schools, split by categories according to the level of Output stream 3 intervention (none, minimum or medium). The data comes from the annual school census from 2009/10 and 2013/14.

Jigawa's schools by level of ESSPIN intervention	None	Minimum	Med. (1)	Med. (2)	
Distance from local government authority headquarters	18.9	16.8	15.2	13.7	-
Age of the school in 2014	29	33.6	36.5	36.3	+
Urban (%)	7.1	13.2	23.6	35	+
Nomadic (%)	10.5	5.5	3.9	0	-
Islamic (%)	5.2	3.7	4.2	7	
Double shift (%)	0.9	1.8	1.4	0.7	
Had parent–teacher association in 2009/10 (%)	95.1	96.2	97.7	99.3	+
Had SBMC in 2009/10 (%)	92.5	94.7	96.3	96.4	+
PTR in 2009/10	47.3	41.3	40.5	37.8	-
PTR in 2013/14	63.7	45.1	49.7	42.4	-
Change in PTR (%)	51.5	35	46.6	39.5	
Number of classrooms in 2009/10	4.4	4.1	4.5	5.4	
Number of teachers in 2009/10	4.4	6.5	9.3	11	+
Primary enrolment in 2009/10	182.9	236.9	350.2	397.9	+
Change in enrolment 2009/10–2013/14 (%)	30.2	28.3	27.5	31.4	
% of teachers with academic diploma/degree	32.3	37.7	42.7	43.2	+
% of teachers with PGDE, BEd or MEd	3	2.3	3.7	5.3	
% of teachers with NCE, Grade II or equivalent	77.7	77.1	75.6	75	
School has a power source (grid/other)	7.2	9.7	14.7	14.2	+
% of classrooms with enough seating	27.8	36.2	34.6	38.9	+
% of classrooms with a good blackboard	39.7	47.2	47	50.9	+
% of classrooms in good condition/minor repairs	49.5	61.3	64.3	65.4	+
School has at least one toilet (%)	17.2	30.2	38.7	35.5	+
Number of schools	1068	437	291	146	

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 Jigawa 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	1,068																	
Minimum	501														6	3	9	
Medium (1)	303											6	3	9	6	3	9	
Medium (2)	198	5*	5*	9*	10*	5*	9*											

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

Annex C ESSPIN Output stream 4 interventions

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

	Level of Output Stream 3 intervention	2010/11			2011/12			2012/13			2013/14			De facto phase
		S	P	M	S	P	M	S	P	M	S	P	M	
Jigawa	None/Minimum													control
	Minimum										2			control
	Medium										7		4	post-CS1
	Medium	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 have benefited from ESSPIN training delivered in 2010 or 2011, so using this information allows us to estimate more accurately how much training the teacher has received.

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 are not significant, allowing us to draw no conclusion as to whether teachers generally improved between CS1 and CS2. The intervention effect is significant only for teachers who were in ESSPIN schools but did not receive the training themselves. Treatment effects for school improvement in Jigawa are negative although not significant, so it is not clear whether teachers who received more training improved faster (or in fact slower) than other teachers between 2012 and 2014.

Table 34. Jigawa: 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	7.3		5.3		5.2		4.9	
	SE	3.8		4.3		4.5		4.5	
Intervention	coefficient	4	*	1.6		3.5		1.1	
	SE	1.4		1.1		2.2		0	
Treatment	coefficient	-3.4		-0.3		-1.7		-0.2	
	SE	1.8		1.5		2.6		1.2	
	N	203		203		197		197	

Note: * indicates a significant coefficient ($p < .05$); adjusted results are adjusted for the length of time a teacher has been in his/her current school – and therefore whether he/she 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). Similarly, no significant difference in the change in headteacher effectiveness was found based on the level of headteacher training received. 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. Jigawa: Headteacher effectiveness difference in differences (regression)

Regression on number of criteria met (/7)		Intervention variable					
		School improvement		Training		Training (adjusted for start date)	
Time (CS2 v. CS1)	coefficient	-0.37		-0.47		-0.39	
	SE	0.43		0.35		0.32	
Intervention	coefficient	0.15		0.04		0.04	
	SE	0.43		0.03		0.03	
Treatment	coefficient	0.12		0.02		-0.01	
	SE	0.55		0.06		0.04	
	N	184		183		179	

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

School development planning

Regression analysis on the number of SDP effectiveness criteria fulfilled finds a significant positive treatment effect, with the level of school development planning improving faster in schools with more recent ESSPIN intervention than in other schools (Table 36). The significant negative intervention effect means the level of school development planning was lower in schools with more ESSPIN intervention.

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

Regression on SDP effectiveness criteria fulfilled (/5)			
Time (CS2 v. CS1)	coefficient	0.11	
	SE	0.1	
Intervention	coefficient	-0.22	*
	SE	0.1	
Treatment	coefficient	1.83	*
	SE	0.49	
	N	199	

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 Jigawa depending on the level of ESSPIN intervention (Table 37). The results suggest that both groups of schools became less inclusive between 2012 and 2014 (as shown by the significant negative coefficient). Although the resulting coefficients are not statistically significant, the positive treatment coefficient gives some suggestion that the schools with more ESSPIN activities worsened less (or more slowly) than other schools over the period.

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

Regression on inclusiveness score			
Time (CS2 v. CS1)	coefficient	-12.5	*
	SE	1.91	
Intervention	coefficient	-0.1	
	SE	3.25	
Treatment	coefficient	2.75	
	SE	3	
	N	204	

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 a positive and significant treatment effect, showing that SBMCs who had received training improved faster than those who did not (and in fact worsened) between 2012 and 2014 (Table 38).

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

Regression on number of SBMC functionality criteria met			
Time (CS2 v. CS1)	coefficient	-0.96	
	SE	0.55	
Intervention	coefficient	-0.59	
	SE	0.55	
Treatment	coefficient	3.51	*
	SE	0.8	
	N	197	

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

The regression on women's inclusiveness of SBMCs also finds that women's inclusiveness improved faster between 2012 and 2014 in schools with more ESSPIN support (Table 39).

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

Regression on number of women's inclusiveness criteria met			
Time (CS2 v. CS1)	coefficient	-0.23	
	SE	0.17	
Intervention	coefficient	0.16	
	SE	0.26	
Treatment	coefficient	1.31	*
	SE	0.38	
	N	187	

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

A regression analysis of difference in differences in SBMC children's inclusiveness finds that schools which received more ESSPIN support improved faster than other schools between 2012 and 2014, according to the positive and significant treatment coefficient (Table 40).

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

Regression on number of children's inclusiveness criteria met			
Time (CS2 v. CS1)	coefficient	-0.15	
	SE	0.13	
Intervention	coefficient	0	
	SE	0.18	
Treatment	coefficient	0.67	*
	SE	0.33	
	N	200	

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

School quality

The regression method for difference in differences in school quality finds there was a significant positive treatment effect (Table 41). This means that the more ESSPIN support schools received, the faster they improved their quality score between 2012 and 2014.

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

Regression on school quality score			
Time (CS2 v. CS1)	coefficient	-0.59	
	SE	2.42	
Intervention	coefficient	-3.92	
	SE	3.89	
Treatment	coefficient	20.65	*
	SE	3.53	
	N	171	

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

Pupil learning

In terms of pupil learning achievement, the regression analysis finds that test scores fell between 2012 and 2014 in all schools, although this fall was significant only for grade 4 numeracy (Table 42). The intervention coefficients are positive, and significant in the grade 4 tests, suggesting that pupils generally scored higher results in ESSPIN schools. The treatment coefficients are negative in three of the four tests, suggesting that ESSPIN schools' results may have improved more slowly than non-ESSPIN schools' pupil test scores; however, the coefficients are not statistically significant.

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

Regression on pupil test score		L2		L4		N2		N4	
Time (CS2 v. CS1)	coefficient	-4.52		-3.16		-4.72		-5.82	*
	SE	4.23		3.32		3.77		2.67	
Intervention	coefficient	9.7		5	*	13.74		5.39	*
	SE	6.8		1.72		7.37		1.31	
Treatment	coefficient	-3.31		0.3		-0.84		-0.45	
	SE	8.07		1.62		7.12		1.23	
	N	199		202		201		202	

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