Oxford Policy Management

CAMBRIDGE EDUCATION

esspin
Education Sector Support Programme in Nigeria

UKaid
from the British people

ESSPIN Composite Survey 3

Kano State Report

Katharina Keck

October 2016

Acknowledgements

We thank the following people for their contributions to this report: Stuart Cameron has provided overall technical leadership and oversight of Composite Survey 3 (CS3), including inputs at various stages of producing this report. CS3 has been managed by Ifeatu Nnodu and Shefali Rai. Femi Adegoke of Oxford Policy Management (OPM) Abuja directed the survey team in Abuja. Ekundayo Arogundade and Kelechi Udoh, and their team of state coordinators and monitoring officers, managed the fieldwork. Babatunde Akano and Joshua Moriyonu led data management. David Megill provided sampling support and survey weights. Katharina Keck and Zara Majeed provided analysis of the data. Reg Allen analysed learner assessments, teacher assessments and the teacher motivation survey. Alia Aghajanian provided analysis of the Annual School Census data. Rachita Daga and Sourovi De contributed analysis of School-Based Management Committee (SBMC) and inclusion indicators. Rachel Outhred and Tanya Lone worked on learner assessment analysis and redesign. Stuart Cameron provided helpful comments on an earlier draft of this report.

From the Education Sector Support Programme in Nigeria (ESSPIN), we acknowledge the contribution of Ron Tuck, who first devised the concept of the Composite Survey; we also acknowledge the various leadership roles undertaken in relation to Composite Surveys 1 and 2 (CS1 and CS2) by Stuart Cameron, Kayode Sanni, John Kay, Lilian Breakell, Simon Thomson, Fatima Aboki, Sandra Graham and Jake Ross. For CS3, Abimbola Osuntusa, Fatima Aboki, Helen Pinnock, Kayode Sanni, Laura McInerney, Manjola Kola, Sandra Graham, Simon Thomson, Jake Ross, Saidi Olalekan and Umar Lawan provided valuable information and support. Allan Findlay and Agbenu Onazi provided information on enrolment and school numbers from the Annual School Census.

The Composite Surveys would not have been conducted without UK aid funding provided by the UK Department for International Development (DFID), including the guidance and support of Barbara Payne, Emily Oldmeadow and Esohe Eigbike.

We are very grateful to the School Support Officers (SSOs), Social Mobilisation Officers (SMOs), and other local government and state staff who acted as data collectors for the Composite Surveys. We are also very grateful to the state government officials who shared information on the context in which ESSPIN has been implemented. Finally, we are extremely grateful to all of the head teachers, teachers and students who kindly took the time to participate in this survey.

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

Oxford Policy Management Limited
Registered in England: 3122495

Level 3, Clarendon House
52 Cornmarket St
Oxford OX1 3HJ
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

The Education Sector Support Programme in Nigeria (ESSPIN) (2008–17) seeks to improve learning outcomes for children of basic education age in six Nigerian states: Enugu, Jigawa, Kaduna, Kano, Kwara, and Lagos. The aims of the ESSPIN Composite Surveys are to assess the effects of ESSPIN's integrated School Improvement Programme (SIP), and to report on the quality of education in the six ESSPIN-supported states. ESSPIN is funded by the UK Department for International Development (DFID) and managed by a consortium led by Cambridge Education. The Composite Surveys have been carried out for ESSPIN by Oxford Policy Management (OPM).

This report presents findings for Kano State from the first, second and third rounds of the ESSPIN Composite Survey (CS1, CS2 and CS3). These took place in 2012, 2014 and 2016, respectively. The survey covered a wide range of indicators at the teacher, head teacher, School-Based Management Committee (SBMC), and pupil levels. The aim is to understand change in schools over time, and whether schools which receive intervention through ESSPIN are working better than those which do not. The main findings are as follows:

Head teacher effectiveness has not changed significantly over time in Kano. Around 17% of head teachers met an overall standard for head teacher effectiveness in 2016. However, there were significant improvements in those aspects of head teacher effectiveness on which the ESSPIN training focused in recent years, which was more limited in scope than the training in other states. In particular, head teachers were significantly more likely to have observed lessons and conducted professional development meetings in 2016 compared to 2012. Most schools received the same level of head teacher training over the past four years, so there is no association to report between the level of ESSPIN intervention and the effectiveness of head teachers in Kano.

School development planning in Kano has improved dramatically since 2012. Around 17% of schools met the standard for effective school development planning in 2016, and schools which have had more years of ESSPIN intervention are much better at school development planning than those which have received fewer years of intervention.

Trends in **inclusiveness** are measured by aspects such as whether the head teacher has taken action on learners' attendance, and whether teachers engage boys and girls equally. For each additional year of intervention received, we estimate that schools meet an additional 0.3 inclusiveness criteria—a small but statistically significant difference. However, fewer schools in Kano met the overall inclusiveness standard in 2016 than in 2012, with only 1.5% of schools fully meeting the standard.

SBMCs in Kano have become more functional, and SBMCs in schools which have had more years of ESSPIN intervention are more functional than those in schools which have received little ESSPIN intervention. SBMCs are more inclusive of women and children compared to 2014, but not compared to 2012. ESSPIN's intervention is associated with more inclusive SBMCs.

Teachers trained through ESSPIN perform much better on English and numeracy content knowledge tests than non-ESSPIN-trained teachers, and are more competent on some of our measures. Over time, there has been no change in average teacher competence during a period in which the coverage of teachers trained has tripled, whilst at the same time complementary state investment has dwindled due to economic and political circumstances. Teachers' scores on the content knowledge tests have worsened significantly over time. Therefore, while teachers' competence has not improved over time, receiving ESSPIN's intervention is associated with better teaching.

Overall **school quality** has improved significantly since 2012, according to our composite measure based on head teacher effectiveness, school development planning, SBMC functionality and teacher competence. The proportion of schools reaching the original quality standard has risen from 3% in 2012 to 17% in 2016, with the majority of that increase occurring since 2014. This means that an estimated 790 additional schools have attained that quality standard in 2016 compared to 2012, containing around 432,000 children who are now learning in a better educational environment by this measure. Schools that have received more years of ESSPIN intervention score higher on measures of school quality but they have not improved significantly faster over time than schools with fewer years of ESSPIN intervention.

Controlling for background characteristics, there are no differences in **learning outcomes** by level of ESSPIN intervention. Over time, grade 4 learning outcomes have not changed significantly while Grade 2 learning outcomes have worsened. However, these results have been observed during a period when total pupil enrolments in Kano have risen by approximately 700,000 pupils[1]. So whilst the average learning outcomes have stood still or declined, the total 'volume' of education delivered has increased substantially. It is likely that newly enrolled children come disproportionately from disadvantaged or conflict-affected backgrounds, implying that the equity in the provision of education may also have increased over time.

Table 1, Table 2, Table 3 and Table 4 summarise the key findings.

Table 1: Kano: Change over time – Key indicators in 2012, 2014, 2016

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
Effective head teacher (%)	17.2	11.8	16.5	-0.7	+4.7
School development planning (%)	3.2	0.3	16.9	**+13.7***	**+16.5***
Inclusive (%)	9.3	3.2	1.5	-7.8	-1.8
Functioning SBMC (%)	19.8	15.9	43.2	**+23.3***	**+27.2***
Competent teachers (%)	67.4	54.8	61.6	-5.8	+6.7
Competent teachers (new measure, %)		9	5.2	n/a	-3.8
Good quality school (%)	3.1	3.3	17.4	**+14.3***	**+14.1***
Good quality school (new measure, %)		0.2	0.2	n/a	+0.0
Grade 2 literacy score	435.9	455.1	423.8	-12.1	**-31.3***
Grade 4 literacy score	438.6	436.7	426.7	-11.9	-10.1
Grade 2 numeracy score	487	444.8	423	**-64.0***	**-21.8***
Grade 4 numeracy score	448.1	442.9	440.6	-7.4	-2.2
Note. * indicates statistical significance ($p < .05$)					

[1] From the 2009/10 baseline of 1,883,472, the total increase by 2013/14 was 861,450: equivalent to absorbing 215,000 more children per year for four years in succession. By 2014/15 the total increase since 2012 eased marginally, to 685,060.

Table 2: Kano: Key indicators in 2016, by years of ESSPIN intervention

	Min. (one year)	Med. (two to three years)	Estimated effect of one year of full intervention
Effective head teacher (%)	16.5	16.1	+2.9
School development planning (%)	13.1	50.9	**+13.1***
Inclusive (%)	18.3	52	+0.1
Functioning SBMC (%)	39.9	71.3	**+18.2***
Good quality school (%)	14.2	45.6	+10.7
Good quality school (new measure, %)	0.2	0	n/a
Grade 2 literacy score	419.3	438	**+13.6***
Grade 4 literacy score	418.8	447.2	+14.4
Grade 2 numeracy score	419.8	432.9	+11.8
Grade 4 numeracy score	436	452.7	+7.5

Note. * indicates statistical significance ($p < .05$). Analysis for functioning SBMC is by Output Stream 4 intervention group; analyses for other indicators are by Output Stream 3 intervention group.

Table 3: Kano: Key indicators on SBMCs and inclusion, by ESSPIN Output Stream 4 intervention

	No inter-vention	Post-CS1	Estimated effect of one year of intervention
School meets standard for functioning SBMC (%)	39.9	71.3	**+18.2***
SBMCs' work on inclusion			
Conducted awareness-raising (%)	66.5	76	+7.7
Addressed exclusion (%)	54.7	66.7	+8.3
Took action for commonly excluded groups (%)	17.6	52.8	**+13.7***
Raised issues of children's exclusion (%)	12.7	38.6	+9.5

Note. * indicates statistical significance ($p < .05$). ESSPIN's Output Stream 4 intervention focuses on improving community participation in school improvement.

Table 4: Kano: Teacher competence and test performance, by ESSPIN training

	Non-ESSPIN-trained	ESSPIN-trained	Difference
Competent teachers (%)	68.5	53.0	**-15.5***
Teacher competence standard (excl. curriculum knowledge, %)	79.2	87.7	**+8.5***
Competent teachers (strict measure, %)	3.7	7.0	+3.2
Teachers' English scale	388.2	425.7	**+37.5***
Teachers' mathematics scale	389.3	434.0	**+44.7***

Note. * indicates statistical significance ($p < .05$)

Table of contents

List of figures, tables and boxes

List of abbreviations

ACLED	Armed Conflict Location & Event Data Project
BEd	Bachelor of Education
CAPI	Computer-assisted personal interview
CBO	Community-based organisation
CS1	Composite Survey 1
CS2	Composite Survey 2
CS3	Composite Survey 3
DFID	Department for International Development (UK)
ESSPIN	Education Sector Support Programme in Nigeria
IRT	Item response theory
L2	Grade 2 literacy test
L4	Grade 4 literacy test
LGA	Local Government Authority
LGEA	Local Government Educational Authority
N2	Grade 2 numeracy test
N4	Grade 4 numeracy test
NCE	National Certificate of Education
NGN	Nigerian Naira
OPM	Oxford Policy Management
PGDE	Post-Graduate Diploma in Education
PTR	Pupil–teacher ratio
QISMB	Qur'anic and Islamiyya School Management Board
SBMC	School-Based Management Committee
SDP	School Development Plan
SE	Standard error
SIP	School Improvement Programme
SSIT	State School Improvement Team
SSO	School Support Officer

SUBEB	State Universal Basic Education Board
TSP	Teaching Skills Programme

1 Introduction

ESSPIN (2008–17) seeks to improve learning outcomes for children of basic education age in six Nigerian states: Enugu, Jigawa, Kaduna, Kano, Kwara, and Lagos. The ESSPIN Composite Surveys seek to assess the effects of ESSPIN's integrated SIP, and to report on the quality of education in the six ESSPIN-supported states. ESSPIN is funded by DFID and is managed by a consortium led by Cambridge Education. The Composite Surveys have been carried out for ESSPIN by OPM.

The first two rounds of the Composite Survey were carried out in 2012 and 2014. The surveys address five output indicators: teacher competence, head teacher effectiveness, school development planning, SBMC functionality and inclusive practices in schools. They also address one outcome indicator, school quality, and one impact indicator, pupil learning achievement. The third round of the Composite Survey (CS3) collected comparable data on these indicators in order to provide information on the extent to which key school-level indicators in the six states have improved during the course of the programme.

This report focuses on the Composite Surveys' findings in Kano State. It presents the key findings from CS3, compares these to the findings of the previous rounds of the survey, and draws out the implications of these findings for ESSPIN's contribution to school-level outputs and outcomes in the state.

1.1 ESSPIN's SIP

ESSPIN aims to bring about better learning outcomes for children of basic education school age in six states, with a range of activities at the state, national, local and school levels. It has four output streams, which focus on:

- strengthening federal government systems;
- increasing the capability of state and local governments as regards the governance and management of schools;
- strengthening the capability of primary schools to provide improved learning outcomes; and
- improving inclusion policies and practices in basic education (ESSPIN, 2013).

Under the third of these outputs, ESSPIN's SIP aims to provide and support the use of structured materials that ensure teachers can deliver quality instruction, to strengthen teachers' own understanding of literacy and numeracy concepts, and to improve academic leadership and school development planning by head teachers (Sanni, 2015). The SIP 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 help maintain and continue improving quality gains. At the same time, many of the same schools have been receiving interventions under the fourth output stream, facilitating community involvement and inclusion through SBMCs.

1.2 ESSPIN in Kano State

ESSPIN has worked with the government in Kano since 2009. A key distinctive feature of ESSPIN's involvement in Kano is the rapid scale-up of the programme from 498 schools in 2012 to 5,482 schools in 2013. While most of the states scaled up ESSPIN activities between 2013 and 2014, the scale-up in Kano was particularly substantial and logistically challenging due to the state's large population and large number of schools. When interpreting the findings of this survey,

it is important to bear in mind that the initial 498 schools differ, in regard to key characteristics, from those schools which were included in the later scale-up.

ESSPIN's support to schools in Kano has encompassed all three elements of the SIP: support to teachers, support to head teachers, and support in relation to school development planning. Teachers have received training on basic teaching skills, the use of classroom organisation, teaching aids and giving praise. Head teachers received training on academic leadership, school planning, the management of teachers, and working with the community. This has been reinforced through regular monitoring and support visits by School Support Officers (SSOs). Since 2013/2014, in direct response to the state government's priorities in a resource-constrained context, ESSPIN's activities in Kano have focused particularly on teaching quality, with the SIP being rebranded as the 'Teaching Skills Programme' (TSP). The focus has been on providing teachers with a full package of teacher training, and ensuring that schools are using the literacy and numeracy lesson plans. Head teacher training has mostly been limited to aspects through which head teachers support teaching quality, particularly lesson observations and professional development meetings. The understanding is that the complementary aspects of school leadership and management training for head teachers will be reinstated to fill out the full school improvement model when resources permit. This factor should be noted when interpreting the results on head teacher effectiveness.

In addition to these interventions, since 2012/2013 schools in Kano have received support under ESSPIN's fourth output stream: improving inclusion policies and practices in basic education. ESSPIN has trained civil society members and government officers from the Department of Social Mobilisation, Social Mobilisation Officers (SMOs), to enable them to train and mentor SBMCs. SBMC members, in turn, have been trained on the roles and responsibilities of SBMCs, school planning and management, communication and leadership, change and relationships management, the participation of women and children in school improvement and education decision-making, resource mobilisation and financial processes, and child protection and participation. This has been complemented by follow-up mentoring and monitoring visits by SMOs.

The extent to which schools have received each of the intervention components has varied from one year to the next (see Annex B and Annex C). In terms of Output Stream 3 intervention (Annex B), 254 pilot schools started off receiving two years of 'full intervention' (leadership training, teacher training and school improvement visits), followed by two years of very little intervention activity. In 2012/2013, a further 255 schools received a year of full intervention, while in 2013/2014 the truncated TSP intervention was scaled up to almost all schools in the state. As a result of priorities shifting towards teaching quality, no leadership training was delivered between 2014 and 2016. Instead, several teachers per school received 20 days of teacher training in 2015/2016, compared to between two to six days of teacher training delivered in the other states.

The Output Stream 4 intervention (Annex C) began in Kano in 2012/2013, and 1,338 schools have received at least one year of full Output Stream 4 intervention. The Output Stream 4 intervention has therefore not yet been fully scaled up in Kano, with the majority of schools (4,910) having received only two days of SBMC training to date. Table 5 shows what proportion of schools in Kano received Output Stream 3 and Output Stream 4 intervention in any given year.

Table 5: Proportion of schools receiving ESSPIN Output Stream 3 and Output Stream 4 interventions

%	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	Any year
Full package of Output Stream 3, including leadership training	4	4	0	4	92	0	0	92
Any Output Stream 3 intervention	4	4	0	4	92	98	100	100
Full package of Output Stream 4	0	0	0	5	20	9	9	25

Note: Proportions are calculated relative to the estimated total number of schools in 2015/16, and so these are not perfectly accurate for other years because the total number of schools changes slightly from year to year. Where Annual School 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 Annual School Census

The expansion of the programme to all schools in the state required a changed model for delivering training, with the training located closer to schools. During the pilot phase of ESSPIN (2009/10 and 2010/11), State School Improvement Teams (SSITs), trained directly by ESSPIN staff, were responsible for supporting, and training head teachers and teachers directly. As the programme expanded, the SSOs were trained by the SSITs and ESSPIN. SSIT members received ongoing training and support over a three-year period to develop the capacity of the State Universal Basic Education Board (SUBEB) Advisory Service Unit, which was established with ESSPIN's support, and to develop the capacity of SSOs, to enable them to lead the school improvement process at school level. Stakeholders suggest that the change in model was well implemented and has not affected the delivery of the SIP.

However, the change in model is likely to have affected the *quality* of implementation to some extent. Programme staff argue that locating training closer to the schools has longer-term benefits, but that in the shorter term the quality standards of the pilot programme might not be fully upheld as the new, much larger numbers of trainers, who typically have lower qualifications than those in the first wave, develop competencies. Because of the large number of schools in Kano, SSITs in Kano train larger groups and have more individuals to support and build relationships with than SSITs in other states.

1.3 Contextual factors and their implications for the SIP in Kano

This section describes key aspects of the backdrop against which ESSPIN's implementation in Kano has taken place in recent years. These contextual factors are relevant when interpreting the changes in school-level outputs and outcomes between CS1, CS2 and CS3. While such changes may have resulted from ESSPIN's support, they may also have been driven by other developments in the state over this period. This section considers:

1. contextual factors or changes that **affect all schools in the state**, whether positively or negatively. These factors may help explain trends across the state as a whole; and

2. contextual factors or changes that **particularly affect schools with more ESSPIN intervention**. These factors may interfere with the analysis of ESSPIN's impact.

Contextual factors that have affected all schools in the state

The most significant contextual change over the last couple of years has been the sharp drop in oil prices, which has had major implications for the fiscal situation in all states. The decline in federal revenues has affected the federal allocations, which are the main source of funding for the SIP. For example, stakeholders report that the second phase of teacher training was affected by political transitions and unstable funding, resulting in irregular monitoring and insufficient coverage of training. The expenditure per child from the budget release for the SIP has decreased sharply in 2015/16 (see Table 6), which has implications for the scale and quality of delivery of the SIP. We generally assume that changes to ESSPIN's intervention only show an effect in one year's time (see Section 2.1.1), so some of the effect of the decline in SIP funding may not yet have been felt by the time of the CS3 survey. In addition to the direct implications for the SIP, there are indirect effects of revenue declines: teachers are not always being paid on time. This could negatively affect teachers' motivation and undermine teacher attendance, head teachers' perceptions of their own ability to influence attendance, and the extent to which teachers apply the new skills gained through ESSPIN training.

Table 6: Kano state primary enrolment and budget release for the SIP

Year	Total enrolment	Budget release (Nigerian Naira (NGN))	Expenditure/child
2013/14	2,743,647	61,000,000	22.2
2014/15	2,876,367	61,000,000	21.2
2015/16	2,985,687	31,000,000	10.4

Source: Kano Annual School Censuses and Kano State Govt Quarterly Monitoring Reports

According to Annual School Census data, enrolment in the state increased dramatically by more than 45% between 2009 and 2013, but stabilised between 2013 and 2014 (Table 7). Between 2009/10 and 2013/14, 861,450 additional children were absorbed by the school system, equivalent to an average of about 215,000 children per year. Enrolment increases are likely to affect both the number and the composition of pupils, as learners from disadvantaged and conflict-affected backgrounds, who would previously have been excluded are presumably among the new entrants to the system[2]. This means that increases in enrolment are likely to come disproportionately from households that are poorest and where the learner's parents have relatively low levels of education, implying that the learners may be less well prepared to enter school.

Stakeholders report that the sharp enrolment increases in recent years have put pressure on the schooling system to supply new classrooms and schools. The addition of an average of 200 new schools per year for five years is likely to have put considerable stress on the schooling system.

Stakeholders suggest that teacher recruitment has kept pace with enrolment increases, although there is some concern about teaching staff being under-qualified and this is likely to have prompted the SIP's shift to focusing on teaching quality. Despite new teachers being recruited, the Annual School Census data show that there has been an increase in pupil–teacher ratios (PTRs) from 63 in 2009 to 69 in 2014. Increases in the PTRs would tend to reduce teachers' abilities to ensure all students achieve good learning outcomes.

In 2014/2015, almost the same number of girls as boys were enrolled in primary schools in Kano, with 49.5% of enrolled pupils being female (1,272,441 female pupils, compared with 1,296,091

[2] Little data are available on the profile of new learners entering schools. However, there have in the past been strong correlations between school attendance and socio-economic status (Mezger, 2016).

male pupils). Out of the six ESSPIN states, enrolment in Kano is the most equal in terms of gender.

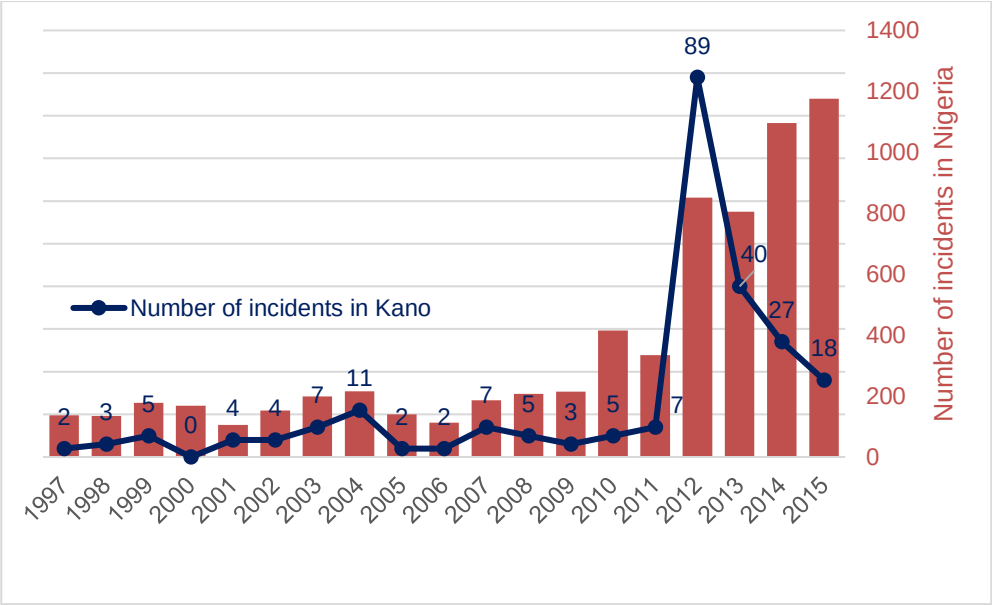
Table 7: Number of schools and enrolment in the 2009, 2013 and 2014 Annual School Censuses

	Enrolment	Number of schools	Enrolment change (%)
2009/10	1,883,472	4,768	
2013/14	2,744,922	5,732	+45.7
2014/15	2,568,532	5,833	-6.4
Overall			+36.4
Note: Enrolment is for Primary Grades 1–6. Source: Annual school census			

There also appear to be some contextual factors that may be supportive of school-level outcomes. Stakeholders report that commitment to the SIP in Kano is strong. There was no change in government in the most recent elections, and stakeholders suggest that the current government is very supportive of ESSPIN's work. Stakeholders report that the government has also been promoting inclusive education and the implementation of compulsory basic education, which should be supportive of ESSPIN's agenda.

While insecurity remains a concern in Kano, the number of incidents of political violence have decreased in recent years. Across Nigeria the number of recorded incidents of political violence and conflict has increased eightfold since 1997, including substantial increases in Kano (Figure 1). The number of incidents of political violence was particularly high in 2011, but the number of both incidents and fatalities has decreased sharply since then, with 18 events and 89 fatalities recorded in 2015 (Table 8). This has coincided with an extended period of heavily militarised security checkpoints within and around Kano City and all major conurbations, which have only latterly been relaxed. Stakeholders report that conflict has not had much impact on ESSPIN activities recently. Localised incidents continue to occur and disrupt ESSPIN activities occasionally, but SSITs appear to have reacted swiftly to any disruptions. Small-scale, qualitative research conducted by ESSPIN in 2014 on the impact of conflict and insecurity reported that children in Kano explained that they continued to feel afraid at school a year after they had experienced a direct armed attack (Coinco, 2014). The psychological effects of armed conflict may well leave traces on children, households and teachers over a longer time span.

Figure 1: Incidents of political violence in Nigeria and Kano



Source: Armed Conflict Location & Event Data Project (ACLED), Version 6 (1997–2015). Note that all events from ACLED are included except for those categorised as protests which did not involve a fatality.

Table 8: Kano: Political violence: Incidents and fatalities, 2010–2015

Variable	2010	2011	2012	2013	2014	2015
Events	5	7	89	40	27	18
Fatalities	6	19	339	186	359	89

Source: ACLED, Version 6 (1997–2015). Note that all events from ACLED are included except for those categorised as protests which did not involve a fatality.

This section highlights two key factors that are likely to dampen the effect of ESSPIN's intervention: first, the decline in federal revenues has affected funding for the SIP, which could not be delivered at full scale. The decline in revenues may also be having an effect on teacher presence and motivation. Secondly, the dramatic increases in enrolment and PTRs are likely to have put substantial pressure on the school system, which may limit the effect of ESSPIN's intervention.

Contextual factors that disproportionately affect schools with more years of ESSPIN intervention

The schools selected for the ESSPIN pilot and those selected for roll-out in 2012/13 differ greatly from those schools which were included only in the later expansion in 2013/14 (see Annex A). The schools with more years of ESSPIN intervention have been established for longer, have better infrastructure in their classrooms, and have more qualified teachers and lower PTRs. They are also much larger and more likely to be located in urban areas. On average, schools with more years of ESSPIN intervention had 1,024 learners, compared to only 397 in the schools that have received fewer years of intervention. Over time, however, schools that have received more years of ESSPIN intervention have also experienced larger increases in enrolment and PTRs than schools that have received fewer years of ESSPIN intervention (see Annex A).

These differences make it difficult to isolate the effect of ESSPIN's intervention. In particular, schools that have received more years of ESSPIN intervention up to now may also have been

more likely to meet ESSPIN standards even before the start of the intervention, due to their background characteristics. Over time, better infrastructure and access to more resources may allow schools with more years of intervention to improve faster, while larger enrolment and PTR increases might limit the rate of improvement. These contextual factors should be kept in mind in the interpretation of the results presented in this report.

2 Methodology and analysis

2.1 Evaluation strategy

2.1.1 ESSPIN intervention groups

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). In practice, the roll-out followed a different implementation plan, with the result that by CS3 all schools in Kano had received some level of ESSPIN SIP activities.

While faster roll-out and greater reach are signs of success for the ESSPIN programme, it presents a difficulty for evaluation as there is now no longer a control group which has received no intervention. The nature, timing and intensity of ESSPIN's intervention varies widely both between and within states (Annex B shows the number of days of leadership training, teacher training and school visits that schools have received under Output Stream 3; Annex C shows the number of days of SBMC training, training in women's and children's participation and mentoring visits received under Output Stream 4).

To simplify the analysis, we focus our analysis on the number of years of full Output Stream 3 intervention that schools have received. Full intervention means that the school has received some leadership training, some teacher training, and some school visits during the year, though the amount of each activity may vary. A further simplification groups the schools into minimum (zero to one years), medium (two to three years), and maximum (four to five years) intervention categories. We also assume that it takes a year for ESSPIN's interventions to have an impact. In line with this, we have not considered activities carried out during 2015/16 when grouping schools into different intervention categories, as the results of these activities will not have emerged by the time of CS3.

While it makes sense to compare the outcomes of schools with different levels of exposure to the intervention, two points must be kept in mind when interpreting the results: (i) there are spill-over effects between schools, which means that staff in minimum intervention group schools might have already been exposed to ESSPIN ideas through informal communication or deliberately by Local Government Education Authority (LGEA) personnel; and (ii) sometimes there are quite extensive changes in school personnel within the state. This means that just because a school has been exposed to the ESSPIN intervention in the past, this does not necessarily mean that its current teachers and head teachers have, and vice versa.

For certain indicators, we alter the classification scheme slightly according to the purpose of our analysis. For example, when examining teacher competence within the CS3 survey, we consider two different groups: teachers who are in schools that have received the ESSPIN intervention but who have not themselves been trained by ESSPIN[3]; and teachers who have been trained by ESSPIN. When examining SBMC functionality and inclusive practices of SBMCs, we classify schools according to the amount of Output Stream 4 intervention received. Schools are classified

[3] Three to six selected teachers within each school attended workshops delivered by SSOs. In some states the same group of teachers continued to receive training year after year, while in other cases attempts were made to spread the training to teachers who had not yet received any. However, teachers in ESSPIN schools are also expected to receive more support through other channels, and particularly through professional development meetings organised by the head teacher (RTI International, 2014; and personal communications from ESSPIN). We distinguish the teachers who received direct training ('ESSPIN-trained') from those who were not themselves directly trained, but are in ESSPIN schools and so are expected to have received support from their head teachers and colleagues ('not-ESSPIN-trained').

as ‘no intervention’ (five or fewer days of Output Stream 4 intervention received), ‘pre-CS1’ (started receiving intervention in 2011/12 or prior to this), and post-CS1 (started receiving intervention in 2012/13 or after).

Learning outcomes – literacy and numeracy in Grades 2 and 4 – are analysed using item response theory (IRT), providing an overall scale of how well children have scored as well as a grouping of children into levels corresponding to the level they are expected to reach by the end of each grade. The distributions of children’s performance across different intervention groups, and across the different survey rounds, are compared. Scores in specific test sub-scales (in literacy these are labelled *receptive*, *fluency in reading and writing*, and *productive*; in numeracy they are *calculation*, *everyday maths*, and *word problems*) are also analysed by intervention group and over time.

Teacher tests – literacy and numeracy – are analysed in a similar manner, also using IRT. As for learners, the analysis includes an overall scale of how well teachers have scored, groupings of teachers according to which grade level they have achieved, and analyses of specific sub-scales within the literacy and numeracy tests.

2.1.2 Types of analysis

The purpose of CS3 is to provide insights into the changes over time in the six states in which ESSPIN works, and to evaluate whether ESSPIN is having an effect in the specific schools in which its school improvement and community inclusion interventions have been applied. We are interested in a wide range of output indicators: teacher competence, head teacher 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 at Grades 2 and 4. For each of these indicators we present in the following chapter two main types of analysis:

1. Change over time between CS1 and CS3, and between CS2 and CS3.

It is important to monitor change over time in how schools function and how much children are learning, both to inform programmes such as ESSPIN and for broader education policy-making. Trends over time in ESSPIN states are likely to reflect both the presence of the intervention and a number of other economic, social and political factors. If ESSPIN has been successful in this aim, then we would expect – other things being equal – that schools in CS3 will have higher output, outcome and impact measures than schools in CS1 and CS2. In practice, however, many other things may not be equal. Changes in enrolment, student profile, state financing, and political commitment may all affect these indicators at the same time. We present these changes over time and, where information is available, consider what may be driving changes aside from ESSPIN intervention.

We use statistical significance tests (t-tests) to give an indication of whether a difference in results between our samples is likely to reflect a genuine difference in the overall populations. Given two ‘populations’ or groups of interest that we wish to compare – say, schools in Kano in 2012 and schools in Kano in 2016 – a common approach is to take a random sample from each group and compare the average performance in one sample to that in the other sample. However, there will be some random variation between the two samples that is due to the set of schools that happened to be sampled. This random variation could result in differences between the two samples even when the two populations are the same. Statistical tests tell us the probability that a difference between the two groups occurred by chance due to random variation in the samples, as opposed to being due to genuine differences in the two populations that the samples were drawn

from. When we are looking at change over time, the t-test tells us the probability that a difference between our 2012 and 2016 sample is due to chance variation between the samples, as opposed to reflecting a genuine change over time in Kano's schools. A probability (sometimes known as the 'p-value') of 5% or less is often taken to be a good threshold for accepting that there is a genuine change, and we mark the result with an asterisk (*) when this is the case.

However, evidence that a change has occurred in Kano's schools does not necessarily imply that this change can be attributed to ESSPIN's activities. There are many other contextual reasons why schools may be improving (or deteriorating) over time, such as the contextual factors that have been described in the previous chapter. When interpreting the reported differences in results between CS1, CS2 and CS3, it is important to bear in mind alternative explanations of why these differences may have arisen (other than through the input of ESSPIN's SIP).

2. Differences between the different levels of intervention categories (minimum, medium and maximum) within the CS3 results. We hypothesise that schools that have received more years of full ESSPIN intervention will have higher output, outcome and impact measures than schools which have received fewer years of intervention.

To test this, we use a continuous measure of the years of full intervention that each school has received (one to five), and calculate the estimated effect of having received one additional year of intervention using a simple regression model. We also show the averages for each intervention group. There are no maximum intervention schools in the Kano sample, so we compare results between the minimum (zero to one years) and medium (two to three years) intervention groups. In terms of Output Stream 4 intervention, there are no 'pre-CS1' schools in the Kano sample, so we compare results between 'no intervention' and 'post-CS1' intervention groups.

Once again, any significant results cannot conclusively be attributed to the ESSPIN intervention. Schools, teachers and pupils who have received fewer years of ESSPIN intervention are likely to differ in regard to background characteristics from those who have received more years of ESSPIN intervention, which is in turn likely to affect their performance. For example, as discussed earlier, schools which have received more years of ESSPIN intervention are typically larger and more urban than schools which were included in the later roll-out (and which have therefore received fewer years of ESSPIN intervention). Controlling for this fully is a more difficult statistical exercise, so we will only attempt this for our outcome and impact measures: school quality and pupil learning outcomes.

For these indicators, we conduct additional analyses in order to understand what basis there might be for making causal attribution of ESSPIN impact. This analysis is described in Sections 5 and 6.

2.2 Sampling, coverage and weights

In CS3, all the schools visited in CS2 were visited again with the intention of collecting data that would enable us to draw inferences about what is happening in the population of schools in Kano, through the use of sample weights. The collected data enable us to analyse changes over time and differences between schools that have received different amounts of ESSPIN intervention.

In Kano the sample size for CS1 was 105 schools. As described in the CS1 report, there was a large amount of variation between schools in the sample, which reduced the precision of the pupil and teacher indicators. To reduce this problem, in CS2 the sample was increased to 175 schools (Megill, 2014). A few of the sampled schools in CS1 and CS2 could not be visited, so the achieved sample was 102 schools for CS1 and 170 schools for CS2.

For CS3, the same schools were visited as during CS2, except for 11 schools which were replaced because they were found to be purely Islamic schools or because the school no longer exists. The number of schools sampled in each of the intervention categories is shown in Table 9.

Table 9: Sample in CS1, CS2, CS3 and population of schools by Output Stream 3 intervention group

Intervention category	Years of full intervention	CS1 sample	CS2 sample	CS3 sample	Population of schools
Minimum	1	56	116	121	4937
Medium	2	11	19	20	255
	3	35	35	34	254
Total		102	170	175	5,482

Note: The sample size shown is the actual sample for which data were collected. Intervention groups reflect the number of years of full Output Stream 3 intervention the schools had received by the end of the 2014/2015 school year.

Within each school, the survey team conducted interviews with the head teacher, the SBMC chairperson or deputy, teachers and pupils. As in CS2, we intended to sample six teachers per school, or all of the eligible teachers in schools with fewer than six teachers. For CS3, we attempted to find the same teachers interviewed during CS2 using their photographs and name information, and to interview them again so as to be able to assess changes over time, as well as rates of change in teacher competence and test results, with more precision. In Kano, we were able to re-sample 47% of teachers sampled during CS2. In schools where we could not re-sample all teachers from CS2, we topped up the sample with eligible teachers selected randomly from the attendance register. This approach to interviewing a panel of teachers allowed us to better track change within *individuals* over time. However, it does imply that we may not have been able to completely capture changes in overall composition that have occurred in the state at the same time. For example, the recent establishment of the Qur'anic and Islamiyya School Management Board (QISMB) has been accompanied by SUBEB's transferral of literacy and numeracy teachers from Islamiyya to mainstream schools. It is therefore not clear whether the composition of our teacher sample might have changed had we selected teachers randomly from sampled schools rather than tracking teachers interviewed during the previous round.

Overall, almost all (over 99%) of the targeted sample of teachers were interviewed and had their lessons observed (Table 10). The administration of the teacher test in CS3 differed to that in CS2. In CS3, teachers were asked to complete the tests in a classroom at the school after lessons had finished; while in CS2, teachers travelled to a testing centre on a separate day where they wrote the tests. The sample coverage for the teacher tests is slightly lower than for the interviews and lesson observations because some teachers did not give their consent to be tested, or were not able to stay for the teacher tests which were held after lessons had finished.

As in CS2, four pupils were sampled per school for each of the tests (Grade 2 literacy, Grade 2 numeracy, Grade 4 literacy and Grade 4 numeracy). Where possible, we sampled pupils who were taught by one of the sampled teachers. The sample coverage of pupils was 97.4%. The number of pupils assessed fell slightly short of the targeted number because some schools had fewer than eight pupils in P2 or P4.

Table 10: Kano: Survey instruments, respondents, sample size and coverage in CS3

	Respondents	Targeted sample size	Number of respondents covered	Sample coverage (% of targeted sample size)
Head teacher interview	Head teachers	175	175	100%
SBMC interview	SBMC chair person	175	175	100%
Teacher interview	Sampled teachers	840	838	99.8%
Literacy test	Sampled teachers	840	809	96%
Numeracy test	Sampled teachers	840	812	97%
Lesson obs	Sampled teachers	840	837	99.6%
L2	Sampled P2 pupils	700	696	99%
N2	Sampled P2 pupils	700	691	99%
L4	Sampled P4 pupils	700	671	96%
N4	Sampled P4 pupils	700	670	96%

Note. In this table and 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. The 'targeted sample size' for teachers represents six teachers per school, or the number of eligible teachers in schools where this is less than six.

2.2.1 Sample weights

Simple averages of the results from the Composite Survey data would not be representative of what is happening across the state, because the profile of schools included in the survey is not identical to the profile of schools in the state as a whole. We address this by applying sample weights which give greater weight to the results obtained from schools that are relatively under-represented in the survey[4]. Sample weights were calculated for the CS1, CS2 and CS3 schools, teachers, and learners. A smoothing technique was also applied to reduce the variability of the weights and to avoid the design effects problem encountered in the CS1 analysis (see Megill, 2014b).

Most of the analysis that follows applies weights to sample statistics calculated within each round and intervention group, which can then be used as estimates of the whole population of schools in the six ESSPIN states. However, part of the analysis compares change within individual schools over time. For this we are limited to the set of schools which were sampled at each of the time points over which the comparison is conducted (e.g. an analysis of change in individual schools between CS2 and CS3 is limited to those schools included in both the CS2 and CS3 rounds of the survey). Additional sets of weights were calculated for use with these 'panels' of schools.

In addition, because we re-sampled teachers from CS2, in some of our analyses we compare how the same teacher performed in CS3 compared to CS2. For this we are limited to the set of teachers who were sampled during both CS2 and CS3. Another set of weights was calculated for use with this panel of teachers.

[4] The CS3 sample comprised 121 schools in the minimum intervention group and 54 schools in the medium intervention group. This under-represents the proportion of minimum intervention schools in Kano State as a whole (91% of schools in Kano are in the minimum intervention group, but our sample proportion is only 69%). To adjust for this, these schools are assigned a larger sampling weight.

2.3 Fieldwork and instruments

Fieldwork for CS3 was conducted using computer-assisted personal interviews (CAPI) during April–June 2016. We made a number of changes to instruments to take on board some additional concerns and to make use of innovations introduced in other recent Nigerian school surveys (described in detail in the CS3 Overall Technical Report). At the same time we retained the questionnaire items required for comparability with previous rounds of the Composite Survey.

Data were collected on teacher competence, head teacher effectiveness, school development planning, inclusive practices in schools, SBMC functionality, teacher competence, teacher subject knowledge and learning outcomes of children in Grades 2 and 4 in English and mathematics. The following activities were carried out as part of the data collection:

- structured interviews with head teachers, SBMC chairpersons and teachers;
- teacher tests in English literacy and numeracy;
- lesson observations; and
- literacy and numeracy tests for pupils in Primary Grades 2 and 4.

The instruments were pre-tested over two days in Abuja during April 2016. State coordinators and monitoring officers collected the data on CAPI after they had been trained on the instruments. Minor revisions were made to the instruments, in consultation with state coordinators.

As in CS2, pupil assessments in CS3 were administered using CAPI. Children were given a printed pupil book to read and write in. The interviewers made use of a tablet computer, which prompted them on the questions the children were to be asked orally, gave instructions on the administration of the different test items, including timing, and allowed them to input whether each part of each question was answered correctly or incorrectly (or not attempted at all) by the pupil. A number of changes were made to the CAPI systems and manuals for the administration of the pupil tests, to make them easier to train on and administer. This included a clear manual with consistent instructions across questions of a particular type, automated timers for timed questions, and translations into Hausa, Igbo and Yoruba of text that did not need to be read in English.

3 School management and head teachers

ESSPIN's interventions include leadership training for head teachers on managing the school and its teachers, planning for the school's development, advocating for more resources, and ensuring that the school is inclusive. ESSPIN also supports the development of SBMCs. This includes training and mentoring on how SBMCs can encourage the participation of women and children. This chapter examines how well schools in Kano are doing on each of these fronts.

ESSPIN's logframe identifies and defines a number of indicators related to school management, inclusiveness and SBMCs. The logframe groups these indicators into a set of 'standards' or composite indicators. These are as follows:

- Head teacher effectiveness: A head teacher is deemed to be effective if they engage in a set of practices that include observing teachers' lessons, holding professional development meetings with teachers, monitoring teacher attendance, keeping records, and ensuring that the school adheres to a regular schedule.
- School development planning: As part of the SIP, schools are encouraged to carry out a self-review process involving the head teacher, teachers, SBMCs, parents and other community members. The aim of this process is to identify the school's strengths and weaknesses, and then list the steps that need to be taken to improve it in a School Development Plan (SDP). The SDP can also be used to request resources from local government or the community. The associated logframe standard assesses whether a self-evaluation has been carried out, whether the school has an SDP, and whether it has implemented the activities in its SDP.
- School inclusiveness: This refers to the extent to which the school makes an effort to include all learners, regardless of gender or socio-economic background. Inclusiveness is assessed on the basis of the steps listed in the SDP and actions taken to boost access, as well as the extent to which teachers encourage the participation of all children in the classroom.
- SBMCs' functionality and performance: The associated standards assess the extent to which SBMCs are functioning and active, and the degree to which they ensure that women and children are actively participating in their activities.

The rest of this section describes each of these standards and then presents associated findings from the Composite Surveys.

3.1 Head teacher effectiveness

Box 1: Head teacher effectiveness: Key findings

- In 2016, 17% of head teachers in our sample meet our standard for an effective head teacher. This does not represent a significant change compared to 2012 or 2014.
- Head teachers in Kano are more likely to have carried out lesson observations and professional development meetings in 2016, compared to 2014 and 2012. These two indicators seem to have been a particular focus of the leadership training in Kano since 2012.
- Head teachers in schools that had received more years of ESSPIN intervention by 2016 are no more effective than head teachers from schools that had received fewer years of ESSPIN intervention. This may in part be explained by the fact that, since CS1, most schools had received the same amount of leadership training (nine days in 2013/2014).

ESSPIN defines head teacher effectiveness by reference to seven criteria set out in its logframe (see Box 2). The first two criteria relate to the pedagogical support that head teachers provide to teachers, the next relates to the steps that head teachers take to boost teacher attendance, and the final four relate to school management practices that have implications for time on task.

Box 2: Logframe criteria for head teacher effectiveness

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

1) have carried out two or more lesson observations in the past two weeks;
2) have held four or more professional development meetings since the start of the school year (NB: the surveys took place more than nine months into the school year);
3) school has a teacher attendance book and the head teacher 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; 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).

Overall, around 17% of head teachers in Kano meet the overall standard for head teacher effectiveness. This does not represent a statistically significant change since 2012 or 2014. It should be noted that head teachers in Kano schools have received only nine days of leadership training (during 2013/2014) since CS1 was carried out, and this training appears to have had a particular focus on lesson observations and professional development meetings. This training was delivered just prior to CS2, and assuming a one-year lag in any intervention effect, we would only notice any effect of this training by CS3.

On these indicators, there appears to have been substantial improvement in 2016 compared to 2014 and 2012. On average, half of the head teachers had carried out two or more lesson observations, compared to under 10% in 2012. Head teachers appear to be getting better at observing their teachers regularly, and at documenting their observations with lesson observation sheets. There has also been rapid progress in the proportion who carried out at least one professional development meeting during the last school term, from 15% in 2012 to 32% in 2016.

Head teachers are significantly less likely to take action on teacher attendance compared to 2012, but equally as likely to take action as they were in 2014. One possible explanation is that they have limited efficacy to influence teacher attendance (partly because they lack the authority to take disciplinary action against teachers). Another is that their motivation has been undermined by the fiscal crisis and associated delays in salary payments.

Fewer schools also conformed to a 35-minute lesson length in 2016 than in 2012 (measured as a lesson length of between 30 and 40 minutes). A length of 35 minutes was formerly considered the standard lesson length across the six states. However, schools have been encouraged to adopt 60-minute lessons, in line with ESSPIN lesson plans, which are intended to be taught over one hour. Longer lessons should therefore arguably be discounted as an indicator of poor head teacher effectiveness, as they may reflect a shift towards one-hour lessons in literacy and numeracy. We have therefore calculated a new indicator, defined as the proportion of schools in which at least half of the observed lessons are at least 30 minutes in length. On this indicator, schools also declined in 2016 compared to 2014, suggesting that lessons are in fact becoming shorter rather than longer. This may partly reflect the effect of observation on teachers. For example, they may be teaching components of lesson plans discussed during training, in a bid to impress the observer, but be unable to work these ideas into a full lesson. Even if this is the case, the short lesson times suggest that teachers have difficulty in planning lesson activities that fill a set duration.

Schools have worsened since 2012 in terms of our indicator of a clear opening time – the extent to which learners and teachers agreed on what the school's opening time is. This appeared to be a result of confusion among learners in particular: on average, only 37% of the learners in each school could agree on its opening time. The reasons for the worsening over time are not clear. However, it can be questioned whether this is a good indicator of school management. Field observations suggested that children were confused over whether to consider the time that they arrived at the school, the time of assembly, or the time when lessons started, as the school opening time.

A better indicator of whether schools maintain a clear schedule are the proportion of teachers and learners present in their classrooms within 30 minutes of the school opening time, an indicator based on direct observation. In 2016, 76% of schools adhered to the indicator that at least half the teachers and learners are in their classrooms on time in the morning. This represents a statistically significant increase compared to 2014, but not compared to 2012. This suggests that teacher and learner attendance has been improving somewhat, although there are still a substantial number of schools in which not even half the teachers are present at the beginning of the day.

Table 11: Kano: Head teacher effectiveness in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
(1) Lesson observations (%)	7.2	25.1	51.8	**+44.7***	**+26.7***
-- No. lesson observations in past two weeks	0.9	1.4	3	**+2.1***	**+1.6***
(2) Professional development meetings (%)	14.9	12.2	32.4	**+17.5***	**+20.1***
-- No. professional development meetings last term	1.5	1	1.1	-0.3	+0.2
(3) Action on teacher attendance (%)	76.5	46.8	50.4	**-26.1***	+3.5
(4) Clear opening time (%)	55.9	30.5	9.6	**-46.3***	**-20.9***
-- Learners who agree on opening time (%)		52.5	36.8	n/a	**-15.7***
-- Teachers who agree on opening time (%)		74.2	67	n/a	**-7.1***
(5) In class on time in morning (%)	65.9	50.5	76	+10.1	**+25.5***
(6) Appropriate morning break (%)	83.5	85.7	87.7	+4.2	+1.9
(7) 35-minute lesson length (%)	31.1	69.5	32.1	+1.1	**-37.4***
-- Lesson length 30 min or longer (%)		84.6	36.1	n/a	**-48.5***
Number of criteria fulfilled (out of seven)	3.6	3	3.3	-0.3	+0.3
Effective head teacher (five out of seven criteria met) (%)	17.2	11.8	16.5	-0.7	+4.7
Note. * indicates statistical significance (p < .05)					

Are the schools that have had more years of intervention under ESSPIN's Output Stream 3 doing better than those which have received the intervention only recently? We present averages across the two intervention groups: those that have only had one year of full intervention up to 2014–15 (minimum), and those that had two to three years (medium) (Table 12). We also estimate the effect of one full year of intervention.

Overall, minimum and medium intervention schools do not differ greatly from one another in terms of head teacher effectiveness. Approximately 16% of head teachers meet the effectiveness standard in each group. This result is perhaps not so surprising, considering that half of the schools in the medium intervention group have received the same level of head teacher intervention as those in the minimum intervention group over the past four years (only nine days of leadership training in 2013/2014). Schools that received additional leadership training did so several years ago, and it is possible that the effects have faded over time. Stakeholders also report that many head teachers from the initial pilot schools have since retired, thereby diluting any potential effect of the training.

Surprisingly, the only statistically significant effect of more ESSPIN intervention is that teachers are *less* likely to agree on the school's opening time – a result which is difficult to explain. On some other indicators, head teachers performed better with more years of ESSPIN intervention, but the results do not reach statistical significance—probably because there is a lot of variability between schools, which makes it more difficult to say with a high level of certainty that a difference is likely to exist in the overall population.

For each year of ESSPIN intervention, head teachers are 12% more likely to have conducted a lesson observation, and 13% more likely to take action on teacher attendance; although neither of these results reach statistical significance. Teachers in schools which have received more years of intervention are 13% less likely to adhere to the standard 35-minute lesson duration with each additional year of intervention, but again this is not statistically significant. It would be worth

exploring why lesson durations appear to be getting shorter, and why this trend seems to be stronger in schools which have received more years of ESSPIN intervention.

Table 12: Kano: Head teacher effectiveness in 2016, by ESSPIN intervention

	Min. (1 year)	Med. (2–3 years)	Estimated effect of one year of full intervention
(1) Lesson observations (%)	49.8	70.7	+11.7
-- No. lesson observations in past two weeks	2.9	4.2	+0.9
(2) Professional development meetings (%)	32.9	27.7	+0.1
-- No. professional development meetings last term	1.1	1.4	+0.3
(3) Action on teacher attendance (%)	48.1	70.7	+13.2
-- School has a teacher attendance book (%)	98.8	98.8	-0.2
(4) Clear opening time (%)	9.9	6.2	-0.7
-- Learners who agree on opening time (%)	36	44.1	+2.9
-- Teachers who agree on opening time (%)	68.2	56.3	**-6.6***
(5) In class on time in morning (%)	74.9	85.3	+6.2
-- Classes where learners present on time (%)	97.1	95.9	-1.7
-- Classes where teachers present on time (%)	65.2	73.7	+4.8
(6) Appropriate morning break (%)	87.6	88.3	-1.1
(7) 35-minute lesson length (%)	34.3	12.8	-12.9
-- Lesson length 30 min or longer (%)	38.4	15.7	-12.4
Number of criteria fulfilled (out of seven)	3.3	3.6	+0.2
Effective head teacher (five out of seven criteria met) (%)	16.5	16.1	+2.9
Additional indicators			
(A1) In class on time after break (%)	84.7	57.3	-7.8
-- Classes where learners present on time (%)	96.9	98.5	+0.7
-- Classes where teachers present on time (%)	70.7	58.9	-3.2
(A2) Teacher absenteeism (%)	23	18.3	-2.1
Note. * indicates statistical significance (p < .05)			

3.2 School development planning

Box 3: School development planning: Key findings

- In 2016, schools had made substantial progress in school development planning, with large and statistically significant increases on the five measured indicators.
- In 2016, we classify 17% of schools as effective at school development planning, compared to only 3% of schools in 2012. Despite the considerable improvement, the proportion of effective schools remains low.
- Schools that have received more years of ESSPIN intervention are much better at school development planning than schools that have received fewer years of ESSPIN intervention, particularly in regard to completing the self-evaluation process, and producing a SDP that meets ESSPIN's criteria.

ESSPIN's leadership training encourages and supports head teachers to review their school's performance and to put together an SDP, which can then be used to advocate for resources from the local government or the community. ESSPIN encourages schools to include a range of measures in their SDPs that go beyond investments in the school's infrastructure and to include other measures to strengthen teaching and learning, and to promote access. Head teachers are also trained on using a cashbook to record the school's expenditure and income. It is expected that these measures will support the effectiveness of school development planning. ESSPIN assesses this on the basis of five criteria, outlined in Box 4:.

Box 4: Logframe criteria for the effectiveness of 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).

In 2016, schools have made substantial progress in school development planning, with large and statistically significant increases on all five measured indicators (

Table 13). There had not been any improvements in the school development indicators in 2014 compared to 2012, so these improvements have occurred primarily over the last two years. As ESSPIN has been rolled out to all schools in Kano, school development planning appears to have become a much more widespread practice. Over 58% of schools have conducted a self-evaluation, and 73% have an SDP available, compared to less than 20% of schools in 2012.

Overall, the average school meets around two of these five criteria, and 17% meet the overall standard for school development planning. While still low, this is a massive improvement on 2012, when the average school met only 0.6 criteria and only 3% of schools met the standard.

Table 13: Kano: SDP effectiveness in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
(1) Written evidence of school self-evaluation process (%)	13.8	9	58.4	+44.6*	+49.3*
(2) SDP available (%)	18.5	7.5	73	+54.5*	+65.5*
(3) SDP contains three or more activities to strengthen teaching and learning (%)	8	1.4	26.6	+18.7*	+25.2*
-- No. activities in SDP to strengthen teaching and learning	0.3	0.1	1.8	+1.4*	+1.6*
(4) Evidence that four or more activities stated in SDP carried out (%)	6.2	0.9	19.2	+13.0*	+18.3*
-- No. activities in SDP carried out	0.3	0.1	1.7	+1.4*	+1.6*
(5) Cashbook up-to-date (%)	15.2	11.7	31.5	+16.3*	+19.8*
-- School has a cashbook (%)	44.6	28.7	81.3	+36.6*	+52.5*
Number of SDP criteria fulfilled (out of five)	0.6	0.3	2.1	+1.5*	+1.8*
School meets effective school development planning standard (four out of five criteria met) (%)	3.2	0.3	16.9	+13.7*	+16.5*
Note. * indicates statistical significance ($p < .05$)					

Schools that have received more years of ESSPIN intervention are much better at school development planning than schools which have received fewer years of ESSPIN intervention, particularly in regard to completing the self-evaluation process, and producing an SDP that meets ESSPIN's criteria. Almost all schools (95%) in the medium intervention category had an SDP available, and 70% of SDPs included at least three activities to strengthen teaching and learning, compared to only 22% of SDPs of schools in the minimum intervention group.

Overall, half the schools in the medium intervention group meet the SDP effectiveness standard, compared to 13% in the minimum intervention group. We estimate that schools meet an additional 0.7 criteria for each year of intervention they receive. As we have noted, this is not a direct estimate of the causal impact of ESSPIN, but it is at least consistent with the idea that longer exposure to ESSPIN's intervention has an impact on school development planning.

Table 14: Kano: SDP effectiveness in 2016, by ESSPIN intervention

	Min.	Med.	Estimated effect of one year of full intervention
(1) Written evidence of school self-evaluation process (%)	55.3	85.7	**+21.4***
(2) SDP available (%)	70.6	95	**+22.3***
(3) SDP contains three or more activities to strengthen teaching and learning (%)	21.8	69.7	**+21.7***
-- No. activities in SDP to strengthen teaching and learning	1.7	2.7	**+0.6***
(4) Evidence that four or more activities stated in SDP carried out (%)	16.2	46.5	+10.5
-- No. activities in SDP carried out	1.6	2.8	**+0.6***
(5) Cashbook up-to-date (%)	32.9	18.3	-6.6
-- School has a cashbook (%)	79.7	95.3	+13.0
Number of SDP criteria fulfilled (out of five)	2	3.2	**+0.7***
School meets effective school development planning standard (%)	13.1	50.9	**+13.1***
Note. * indicates statistical significance (p < .05)			

3.3 School inclusiveness

Box 5: School inclusiveness: Key findings

- Schools in Kano have become less inclusive since 2012, according to our indicator of school inclusiveness.
- In 2016, only 1.5% of schools fully meet our standard for an inclusive school, compared to 9% in 2012.
- Schools that have received more years of ESSPIN intervention are more inclusive than schools that have received fewer years of ESSPIN intervention. For each additional year of intervention received, we estimate that schools meet an additional 0.3 inclusiveness criteria: a small but statistically significant difference.

The criteria on school inclusiveness measure the extent to which the school makes efforts to include all learners, including those from disadvantaged backgrounds. ESSPIN's overall standard for school inclusiveness is based on four criteria (

Box 6: Standard for school inclusiveness

). Further detail on these is provided in the companion Gender and Inclusion Report.

Box 6: Standard for school inclusiveness

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) head teacher 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 in between 60% and 40% of total engagements meet the criterion).

Schools in Kano have become less inclusive compared to 2012 (

Table 15). Only 1.5% of schools are able to fully meet the inclusiveness standard in 2016, compared to 9% of schools in 2012. The proportion of schools partially meeting the inclusiveness standard has declined sharply and significantly since 2012. Compared to 2014, the proportion of schools fully or partially meeting the inclusiveness standard has declined slightly.

The decline in school inclusiveness seems to be driven by head teachers taking less action on learner attendance, and fewer teachers using two or more assessment methods to assess their learners. In 2016, only 36% of teachers reported using at least two different methods, while this was 64% in 2012 and 59% in 2014.

One indicator which has improved since 2012 and 2014 is the proportion of schools who include at least two activities on improving access in their SDP. This may reflect general improvements in school development planning, as noted in the previous section.

Table 15: Kano: School inclusiveness in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
(1) Three or more actions on learner attendance (%)	62.8	22.4	28.8	**-34.0***	+6.4
-- Number of actions on learner attendance	2.7	1.9	2	**-0.7***	+0.1
(2) Two or more activities in SDP on access (%)	4.4	1.9	9.6	+5.2	**+7.7***
-- Number of activities on access	0.1	0.1	0.4	**+0.3***	**+0.4***
(3) >50% of teachers use two or more assessment methods (%)	63.7	59.1	35.8	**-27.8***	**-23.3***
(4) >50% of teachers spatially inclusive and >50% are gender inclusive (%)	26.3	15	11.8	-14.5	-3.3
Number of inclusiveness criteria fulfilled (out of four)	1.6	1	0.9	**-0.7***	-0.1
Inclusiveness score (%)	50.1	38.9	31.2	**-18.9***	**-7.7***
School fully met standard (three to four criteria) (%)	9.3	3.2	1.5	-7.8	-1.8
School partially met standard (two to four criteria) (%)	63.4	31.4	21.7	**-41.7***	-9.7

Note. * indicates statistical significance (p < .05). The inclusiveness score is a total ranging from 0 to 100 and is calculated as follows: $\mathbf{20}(\frac{s_1}{7} + \mathbf{min}\left(\mathbf{1}, \frac{s_2}{5}\right) + \frac{s_3}{3} + \frac{s_4}{6} + \boldsymbol{s_5})$, where s_1 is the number of actions to improve attendance; s_2 is the number of activities in the SDP to improve access for disadvantaged children; s_3 is the average number of assessment methods used by sampled teachers; s_4 is the average number of classroom zones participating in the lesson during lesson observations, and s_5 is the gender equity score (see below).

While inclusive practices in schools have declined over time, schools that have received more years of ESSPIN intervention are more inclusive than schools that have received fewer years of ESSPIN intervention (

Table 16). For each year of intervention, we estimate that schools meet an additional 0.3 inclusiveness criteria – a small but statistically significant difference. Compared to only 18% in the minimum intervention group, 52% of schools in the medium intervention group partially meet the inclusiveness standard. This appears to be quite a large difference, and the estimated effect of an additional year of intervention almost reaches statistical significance ($p = .06$).

Schools that have received more years of ESSPIN intervention are more likely to include at least two activities on access. The other indicators generally point in the expected direction (the estimated effect is positive), but are not statistically significant.

Table 16: Kano: School inclusiveness in 2016, by ESSPIN intervention

	Min.	Med.	Estimated effect of one year of full intervention
(1) Three or more actions on learner attendance (%)	26.1	53.4	+11.1
-- Number of actions on learner attendance	2	2.2	+0.1
(2) Two or more activities in SDP on access (%)	8.5	19.7	**+5.5***
-- Number of activities on access	0.4	0.9	**+0.3***
(3) >50% of teachers use two or more assessment methods (%)	33.1	60.1	+12.8
(4) >50% of teachers spatially inclusive and >50% are gender inclusive (%)	12.1	9.1	-0.2
Number of inclusiveness criteria fulfilled (/4)	0.8	1.4	**+0.3***
Weighted sum inclusiveness score	30.7	35.8	**+2.6***
School fully met standard (three to four criteria) (%)	1.1	5.1	+1.5
School partially met standard (two to four criteria) (%)	18.3	52	+12.2
Additional indicators			
Enrolment increased since last year (%)	58.8	75.8	+10.9
Change in enrolment since last year	0	0.3	+0.1
Note. * indicates statistical significance (p < .05)			

3.4 SBMCs

Box 7: SBMCs: Key findings

- In 2016, SBMCs in Kano are more much functional than SBMCs in 2012 and 2014. About 43% of schools meet our standard of SBMC functionality in 2016, compared to 20% in 2012.
- ESSPIN's Output Stream 4 intervention is associated with better functioning SBMCs. SBMCs meet an estimated additional 1.1 criteria per year of ESSPIN Output Stream 4 intervention.
- Participation of women and children within SBMCs has not changed since 2012, keeping in mind that over two-thirds of the sample have not received training or support in these areas.
- SBMCs in schools that have received ESSPIN's Output Stream 4 intervention are more inclusive of women and children than SBMCs in schools that have received no intervention.

SBMCs are considered by ESSPIN to be functioning well if they meet regularly and work with the community, traditional and religious institutions, and local government to address the school's needs, raise resources for the school, and find ways to tackle exclusion. They are expected to have a women's committee and a children's committee, and to keep financial records. They are also expected to play a supervisory role, marked by regular visits to the school by the chairperson and other SBMC members. In line with this, ESSPIN uses nine criteria to assess SBMC functionality (see

Box 8: Logframe criteria for SBMC functionality

). SBMCs are considered to be effective if they meet at least five of the nine criteria. In most cases, these require evidence to be presented, rather than just accepting the word of the respondent (usually the SBMC chairperson). Thus, they reflect the ability of the SBMC to keep good records of their activities, as well as actually undertaking the activities themselves.

Box 8: Logframe criteria 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:[5]

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

In 2016, SBMCs in Kano are much more functional compared to 2012 and 2014. By 2016, all sampled schools in Kano have an SBMC, compared to only 66% of schools in 2012. In addition, 43% of schools meet the standard for a functioning SBMC, meeting an average of 4.2 out of nine criteria. This compares to 20% of schools meeting the standard in 2012, with an average of two of nine criteria met.

There have been large and significant increases since 2012 and/or 2014 in the proportion of SBMCs that met at least twice in the current school year, networked with other institutions, interacted with LGEAs, and contributed resources for the school. Many more SBMCs now have women's and children's committees, and the proportion of SBMCs whose chairperson had visited the school has increased compared to 2014 (although not compared to 2012). Despite the improvements over time, the proportion of schools that have women's and children's committees and have met at least twice remains low.

SBMCs appear increasingly to be taking actions to improve access to learning for all children. In 2016, a much higher proportion of SBMCs have conducted awareness-raising and have addressed exclusion, compared to 2012[6]. However, a significantly lower proportion of SBMCs are taking action for commonly excluded groups, such as children with disabilities, girls or nomadic children, compared to 2014, possibly because such action is less needed. For example, Kano was the ESSPIN state with the greatest gender equity in enrolment in 2014/2015, and the rapid enrolment increases suggest that many previously excluded children may now be attending school.

[5] 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.
[6] The question posed to SBMCs was whether they had addressed issues that prevent children from attending school or which cause drop-out in the current school year.

Table 17: Kano: SBMC functionality in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	2012–16	2014–16
(1) Two or more meetings this school year (%)	37.2	16.3	46.1	+8.9	**+29.8***
(2) Conducted awareness-raising (%)	32.9	53.2	67.5	**+34.6***	+14.4
(3) Addressed exclusion (%)	24.3	40.3	56	**+31.7***	**+15.7***
(4) Networked with CBOs/institutions/other SBMCs (%)	8.8	56.9	95.8	**+87.0***	**+38.9***
(5) Interacted with LGEA (%)	16.3	11.6	38.6	**+22.3***	**+27.0***
(6) Has a women's committee (%)	2.8	11	34.5	**+31.7***	**+23.4***
(7) Has a children's committee (%)	17	10.6	23.5	+6.5	**+12.9***
(8) Contributed resources for school (%)	38.2	55.1	64.5	**+26.3***	+9.4
(9) Chair visited school three or more times (%)	21.3	6.4	17.1	-4.3	**+10.7***
Number of SBMC functionality criteria met (out of nine)	2	2.9	4.2	**+2.2***	**+1.3***
School meets standard for functioning SBMC (%)	19.8	15.9	43.2	**+23.3***	**+27.2***
Additional indicators: inclusion and drop-out					
(A1) Took action for commonly excluded groups (%)	15.7	37.5	21.3	+5.6	**-16.2***
(A2) Raised issues of children's exclusion (%)	7.4	18.3	15.4	+8.0	-2.9
Additional indicators: organising and mobilising resources					
(A7) School has an SBMC (%)	65.9	90.3	100	**+34.1***	**+9.7***
(A8) Cashbook available (%)	29.9	21.7	65.9	**+36.0***	**+44.2***
(A9) Requested support from LGEA or SUBEB (%)			62.5	n/a	n/a
(A10) Raised cash to support school improvement (%)	19.6	26.3	25.7	+6.1	-0.6
(A11) Mobilised non-cash resources (%)	32.6	51	56.8	**+24.2***	+5.7
(A12) Involved in making SDP (%)		7.8	43.4	n/a	**+35.6***
Note. * indicates statistical significance (p < .05)					

Box 9: Asking SBMCs about inclusion and exclusion

A number of different criteria aim to measure the SBMC's inclusiveness and the actions it has taken on excluded children. These were based on the following questions addressed to the SBMC chairperson. As elsewhere, questions were asked in the local language, with instructions to use a language that the respondent could understand, but not to provide additional explanation or prompts.

Criterion	**Question asked** (with data collector instructions in blue)	Criterion met if…
(2) Conducted awareness-raising	Did the SBMC do anything to raise awareness about the value of education for all boys and girls in the community in the current school year?	Respondent answers yes and can present oral or written evidence

(3) Addressed exclusion	Did the SBMC do anything to address issues which prevent children from attending school or which cause drop-out in the current school year?	Respondent answers yes and can present oral or written evidence
(A1) Took action for commonly excluded groups	Did the SBMC do anything to **support commonly excluded groups** in the **current school year**? You can explain that commonly excluded groups could be orphans, nomadic children, girls, children with disability, ethnic or religious minorities, etc.	Respondent answers yes and can present oral or written evidence
(A2) Raised issues of children's exclusion	Did the SBMC **raise issues of children's exclusion** from school in the community, with the LGEA, or with the state government, in the **current school year**?	Respondent answers yes and can present oral or written evidence
(A3) Raised cash to support vulnerable children	Did the SBMC mobilise any cash to support vulnerable children in the current school year?	Respondent answers yes (no evidence required)
(A4) Monitored drop-out or non-attendance (A5) Communicated with school or community about drop-out (A6) Number of actions taken to address non-attendance	**What actions were taken** to address issues which prevent children from attending school or which cause drop-out in the **current school year**? Do not prompt. This is a multiple response question **– SELECT ALL THAT APPLY** • Monitoring drop-out • Monitoring non-attendance • Communicating with school about drop-out • Communicating with community about drop-out • Other (specify) • Don't know / refused	Respondent answers yes to a previous question (asking whether any action was taken to address these issues) and then provides this information in the follow-up question on what type of action and how many actions were taken. No specific evidence is required

ESSPIN's Output Stream 4 intervention is associated with better functioning SBMCs. On the individual criteria, results are in the expected direction for eight out of nine criteria (the chair visiting the school three or more times being the exception), but are mostly non-significant. However, SBMCs in schools which have had greater exposure to ESSPIN intervention are significantly more likely to have women's and children's committees, and to have interacted with the LGEA.

SBMCs in schools which have had greater exposure to ESSPIN intervention are more likely to be taking actions to improve access for the inclusion of all children (see Box 9 above). In 2016, a

significantly higher proportion of SBMCs had taken action for commonly excluded groups (such as orphans, girls, disabled children) and to have raised money to support vulnerable children. Most schools (and more in the medium intervention group) had communicated with the community about drop-out.

Looking at the overall functionality standard, there is a large and significant difference in the proportion of schools that meet the standard: 71% of schools in the post-CS1 intervention group meet the standard, compared to only 39% of schools that have received less than five days of intervention so far ('no intervention'). We estimate that a year of full Output Stream 4 intervention is associated with 1.1 additional criteria being met.

Table 18: Kano: SBMC functionality in 2016, by ESSPIN intervention

	No intervention	Post-CS1	Estimated effect of one year of full intervention
(1) Two or more meetings this school year (%)	43.8	65.8	+13.2
(2) Conducted awareness-raising (%)	66.5	76	+7.7
(3) Addressed exclusion (%)	54.7	66.7	+8.3
(4) Networked with CBOs/institutions/other SBMCs (%)	95.6	97.3	+0.9
(5) Interacted with LGEA (%)	35.4	66	**+13.9***
(6) Has a women's committee (%)	29.3	78.3	**+25.0***
(7) Has a children's committee (%)	18	70	**+19.6***
(8) Contributed resources for school (%)	62.6	80.2	+9.7
(9) Chair visited school three or more times (%)	17.9	10.6	-4.1
Number of SBMC functionality criteria met (out of nine)	4	6	**+1.1***
School meets standard for functioning SBMC (%)	39.9	71.3	**+18.2***
Additional indicators: inclusion and drop-out			
(A1) Took action for commonly excluded groups (%)	17.6	52.8	**+13.7***
(A2) Raised issues of children's exclusion (%)	12.7	38.6	+9.5
(A3) Raised cash to support vulnerable children (%)	40.7	70.6	**+17.4***
(A4) Monitored drop-out or non-attendance (%)	62.8	71.1	+6.6
(A5) Communicated with school or community about drop-out (%)	83	91.7	+5.3
(A6) No. actions taken to address non-attendance	1.5	2.5	+0.5
Additional indicators: organising and mobilising resources			
(A7) School has an SBMC (%)	100	100	n/a
(A8) Cashbook available (%)	63.1	89.6	**+17.9***
(A9) Requested support from LGEA or SUBEB (%)	59.7	86.5	**+16.8***
(A10) Raised cash to support school improvement (%)	25.9	23.9	-3.2
(A11) Mobilised non-cash resources (%)	54.1	79	+13.6
(A12) Involved in making SDP (%)	40.9	64.8	+13.8
Note. * indicates statistical significance ($p < .05$)			

3.4.1 How inclusive are SBMCs of women and children?

As noted above, SBMCs are expected to have women's and children's committees. We also record a number of other measures of the extent to which SBMCs are inclusive of women's and children's concerns. In each case, there are four criteria and an overall standard (Box 10).

Box 10: Logframe criteria for SBMCs' inclusiveness of women and children

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.[7]

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.[8]

In the CS2 report we noted that there had been a large, significant fall in women's and children's inclusiveness in Kano, with only 1.4% of schools meeting the women's inclusiveness standard and 0.2% of schools meeting the children's inclusiveness standard. By 2016, this trend appears to have reversed. On almost all indicators, there are large and significant improvements compared to 2014, but no change compared to 2012 (Table 19). We do find significant improvements compared to 2012 on two indicators: the proportion of children's committees which have met in the current school year; and the average number of criteria on children's inclusion fulfilled.

Levels of inclusiveness were low in 2012 and thus they continue to be low in 2016. On average, schools meet less than one of the women's and children's inclusiveness criteria. Only 12% of schools meet the standard for women's inclusiveness and 7% of schools meet the standard for children's inclusiveness. The results point towards severe shortfalls remaining in regard to these indicators. By 2016, over two-thirds of the sample of schools in Kano had not received any training on women's and children's participation, or any mentoring visits, and so a lack of progress on these indicators is perhaps not surprising.

Table 19: Kano: SBMCs' inclusion of women and children in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012-16	Change 2014-16
Participation of women in SBMC					

[7] This criterion has been slightly altered since CS1, where it also required that the women's committee have a female leader.
[8] 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.

(1) At least one woman attended two or more meetings (%)	18.3	3.2	22.6	+4.4	**+19.4***
(2) Female member raised an issue (%)	28.4	6.7	38.6	+10.2	**+31.9***
(3) Issue raised by female member led to action (%)	22.1	0.4	18.3	-3.9	**+17.9***
(4) Women's committee met (%)	14.4	8	24.8	+10.4	**+16.8***
No. criteria met (out of four)	0.5	0.2	0.8	+0.3	**+0.7***
Meets standard (three out of four criteria) (%)	20.9	1.4	11.8	-9.0	**+10.4***
Participation of children in SBMC					
(1) At least one child attended two or more meetings (%)	13	1.5	18.7	+5.6	**+17.2***
(2) A child raised an issue (%)	12.9	4.5	22	+9.1	**+17.5***
(3) Issue raised by child led to action (%)	15.8	0.1	15.3	-0.5	**+15.2***
(4) Children's committee met (%)	1.4	2.9	12.6	**+11.2***	**+9.6***
No. criteria met (out of four)	0.3	0.1	0.6	**+0.3***	**+0.5***
Meets standard (three out of four criteria) (%)	7.7	0.2	7.3	-0.4	**+7.0***
Note. * indicates statistical significance (p < .05)					

SBMCs in schools that have received ESSPIN's Output Stream 4 intervention are more inclusive of women and children's concerns than SBMCs in schools that have received no intervention (

Table 20). On all but one indicator, the differences are large and statistically significant. Only 7% of no intervention schools meet the overall standard for women's participation, compared to 55% of post-CS1 schools. Similarly for the children's standard, 44% of post-CS1 schools meet the standard, compared to only 3% of no intervention schools. These findings are consistent with an interpretation that poor improvements over time are driven by a lack of intervention for the majority of schools; and that receiving ESSPIN's Output Stream 4 intervention is in fact associated with large improvements in women's and children's participation in SBMCs.

Table 20: Kano: SBMCs' inclusion of women and children in 2016, by ESSPIN intervention

	No intervention	Post-CS1	Estimated effect of one year of full intervention
Participation of women in SBMC			
(1) At least one woman attended two or more meetings (%)	18.7	56.4	**+14.9***
(2) Female member raised an issue (%)	30.4	81.3	**+23.8***
(3) Issue raised by female member led to action (%)	13.6	58.1	**+14.8***
(4) Women's committee met (%)	19.4	70.3	**+20.0***
No. criteria met (out of four)	0.6	2.4	**+1.0***
Meets standard (three out of four criteria) (%)	6.9	54.6	**+11.4***
Participation of children in SBMC			
(1) At least one child attended two or more meetings (%)	18.1	23.2	+2.9
(2) A child raised an issue (%)	16.2	62.6	**+15.5***
(3) Issue raised by child led to action (%)	11.4	49.5	**+12.1***
(4) Children's committee met (%)	7	59.4	**+12.2***
No. criteria met (out of four)	0.5	1.8	**+0.7***
Meets standard (three out of four criteria) (%)	3.1	44.3	**+7.6***

3.5 Summary: School management and head teachers

Box 11: Kano: School planning and SBMC functionality are improving over time

Comparison of school management indicators between CS1 (2012), CS2 (2014) and CS3 (2016)

shows the average progress in the school management standards in Kano primary schools in CS1, CS2 and CS3. Between 2012 and 2016, schools have improved substantially on their effectiveness of school development planning, and SBMCs have become much more functional. On the other hand, head teacher effectiveness has not improved over time, schools have become less inclusive and participation of women and children in SBMCs has not changed compared to 2012.

Limited change in the participation of women and children in SBMCs is not surprising, given that many schools have not yet received training in these areas. Similarly, the lack of change in head teacher effectiveness may reflect the fact that the SIP has not placed strong emphasis on leadership training in recent years. Areas of leadership training that have received attention (lesson observations and professional development meetings) have improved over time.

Despite the improvements over time on some indicators, the proportion of schools meeting the school management standards remains fairly low, and leaves room for further improvement. In 2016, 17% of schools meet the standard on head teacher effectiveness; 17% on school development planning; 1.5% on inclusion; and 43% on functional SBMCs.

The very low proportion of schools meeting the inclusiveness standard, and the decrease in schools meeting the standard over time, presents a cause for concern and requires further investigation. In particular, it highlights a need to engage with why teachers are less likely to use multiple assessment methods, and to be spatially and gender inclusive. One possibility is that increases in class sizes have made it increasingly difficult for teachers to focus on inclusive teaching practices.

Box 11: Kano: School planning and SBMC functionality are improving over time

Comparison of school management indicators between CS1 (2012), CS2 (2014) and CS3 (2016)

Schools which have received more years of ESSPIN intervention generally perform better than schools which have received fewer years of ESSPIN intervention. In

Box 12: Kano: Schools which have received more years of ESSPIN intervention are better at school management

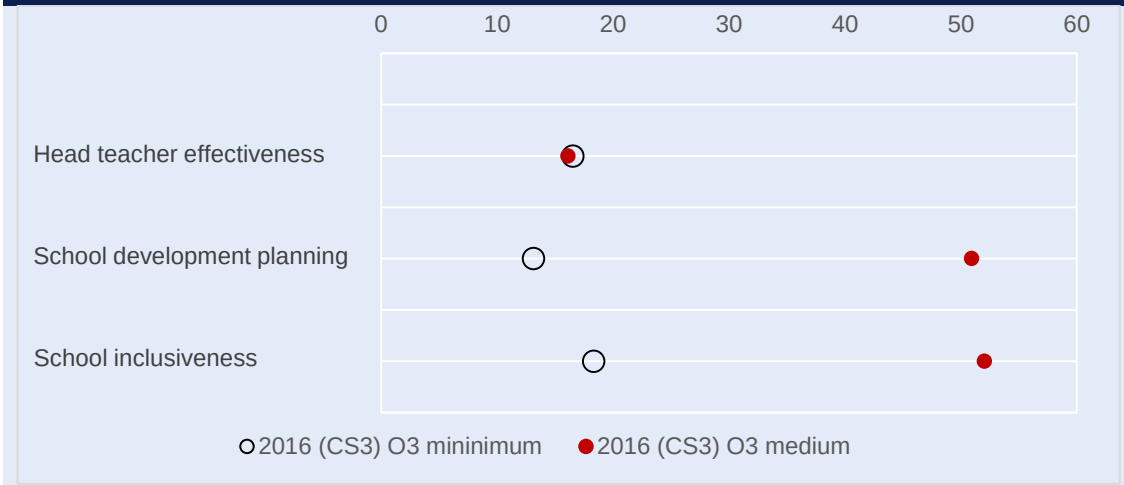
Comparison of school management by Output Stream 3 intervention group in 2016

we can see that a higher proportion of the schools that have received two or three years of Output Stream 3 intervention by 2016 (med.) meet the school development planning and school inclusiveness standards than schools that have received only one year of intervention (min.). Schools which have received more Output Stream 4 intervention are more likely to have well-functioning SBMCs, and are more likely to include women and children in these SBMCs (Box 12). Head teacher effectiveness is the only standard on which schools which have received more year years of ESSPIN intervention do not perform better than schools which have received fewer years of ESSPIN intervention, although this can be partly explained by the fact that the great majority of schools have received the same volume of leadership training in the last four years.

While we cannot attribute these results to ESSPIN's intervention directly, they are consistent with the hypothesis that ESSPIN's intervention is contributing to better school management. In the medium intervention group of Output Stream 3 intervention, between 16% (head teacher effectiveness) and 52% (school inclusiveness) meet each standard. In the post-CS1 intervention group for Output Stream 4 intervention, between 44% (children's inclusiveness) and 71% (SBMC functionality) meet the standards. On most indicators, this represents a large difference compared to the minimum / no intervention group. However, it also means that with several years of full ESSPIN intervention, a substantial number of schools are not yet able to meet these indicators.

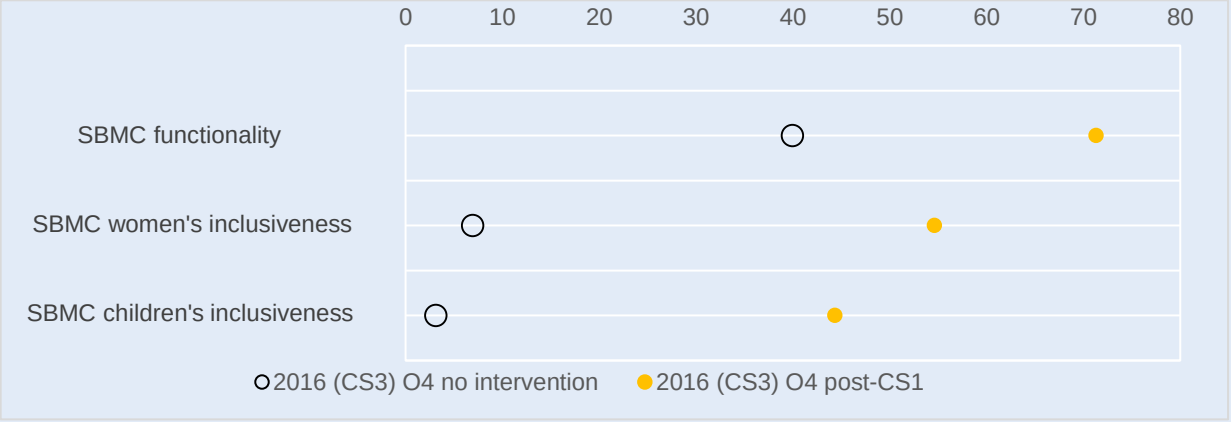
Box 12: Kano: Schools which have received more years of ESSPIN intervention are better at school management

Comparison of school management by Output Stream 3 intervention group in 2016



Box 13: Kano: SBMCs which have received more years of ESSPIN intervention are more functional and inclusive

Comparison of school management by Output Stream 4 intervention group in 2016



4 Teachers

The SIP includes training for teachers on basic literacy and numeracy, and teaching skills (the use of teaching aids, participation and giving praise, and techniques for classroom organisation). This chapter first looks at changes in teacher competence as measured by ESSPIN's logframe indicator. It then looks in more detail at teachers' performance on a set of literacy and numeracy content knowledge tests. Finally, it looks at changes in teacher motivation (which may have important implications for the extent to which changes in teacher competence translate into improvements in teaching practice).

Box 14: Teachers: Key findings

- We find no change in teacher competence between 2012 and 2016.
- In 2016, 62% of teachers are competent on a measure based on lesson observations and teachers' curriculum knowledge. However, only 5% of teachers meet a stricter standard of competence that includes scoring over 50% in a literacy and numeracy test.
- Teachers trained by ESSPIN are performing slightly better than non-ESSPIN-trained teachers. In particular, they are more likely to use teaching aids interactively, and to assign individual or group tasks during their lessons.
- In 2016, teachers perform significantly worse across all domains of the English literacy and numeracy tests, compared to 2014.
- Teachers who had received ESSPIN training performed better on the content knowledge tests in 2016, compared to non-ESSPIN-trained teachers.
- Teachers who had received ESSPIN training in 2014 had similar levels of motivation in 2016, compared to teachers who had not received ESSPIN training in 2014.

4.1 Teacher competence

The ESSPIN logframe sets four criteria for judging the competence of teachers: one relates to curriculum knowledge (although this applies only to teachers who teach English or mathematics), two relate to teaching practices, and one relates to classroom organisation. Teachers are defined as competent if they meet three of the four criteria (two of the three relevant criteria in the case of those who do not teach English or mathematics; see Box 15).

For CS2 and CS3, a stricter version of the competence indicator was developed. The criterion on using at least one teaching aid during the lesson observation was changed to exclude reading from, writing on, or having pupils copy from, the blackboard, since this is considered to be a poor use of a teaching aid that is less likely to enhance learning. In addition, a fifth criterion was added that is based on teachers' performance in content knowledge tests. 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 15: Criteria 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 mathematics. 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 and CS3, stricter criteria for teacher competence were introduced. These modified (2) to exclude reading from or writing on, or having pupils copy from, the blackboard as a use of a teaching aid. A fifth criterion was added:

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

Teachers in Kano schools are showing significant improvements on two of the teacher competence indicators. In 2016, teachers are more likely to be using more praise than reprimands during their lessons than they were in 2012. Teachers are also significantly more likely to use at least one teaching aid during their lesson, with almost all teachers doing so (98.7%), compared to 86% in 2012. However, reading from, writing on, or having pupils copy from, the blackboard is the only teaching aid used by many teachers (42%), who fail to make use of more innovative teaching aids.

On the other hand, the proportion of teachers who assign at least two individual or group tasks has decreased dramatically since 2012, from 60% to 19%. The high proportion of teachers apparently assigning individual and group tasks in CS1 may include some measurement error. In CS3, for example, it was made clear that the definition of an individual task should not include learners copying from the blackboard – an activity that might be interpreted as an individual task without this clear guidance. Still, the low proportion of teachers who are assigning these tasks in 2016 leaves much room for improvement.

Performance in the content knowledge tests was particularly poor in Kano, and worsened significantly between 2014 and 2016. Only 11% of teachers were able to score at least 50% in both the English literacy and numeracy test, compared to 19% in 2014. It should be noted that 30% of the sampled schools in Kano are Islamiyya schools, which typically do not have literacy and numeracy teachers. As a result, it is expected that these teachers would perform more poorly in literacy and numeracy content knowledge tests, and this is likely to be a contributing factor to poor test performance overall. However, since the composition of schools in the sample has not changed since 2014, this does not help to explain the drop in test performance over time. Further analysis of the content knowledge tests is presented in the section that follows.

Overall, teacher competence has not changed significantly compared to 2012 or 2014, by reference to either version of the teacher competence standard. About 62% of teachers meet the original version of the competence standard in 2016, but only 5% meet the stricter version of the standard. This is being driven by the very poor performance in the content knowledge tests.

Table 21: Kano: Teacher competence in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
(1) Knowledge of Eng./mathematics curriculum (%)	41.9	24.6	29.1	-12.8	+4.5
(2a) Use of one or more teaching aid (%)	85.5	93.2	98.7	+13.2*	+5.5*
(2b) Use of one or more teaching aid, excl. blackboard (%)		52.1	56.4	n/a	+4.3
(3) Praise more than reprimand (%)	63.5	70.9	82.0	+18.4*	+11.0*
(4) Assigns two or more ind./group tasks (%)	59.5	26.7	19.1	-40.4*	-7.5
(5) Passes English and mathematics test		18.9	10.7	n/a	-8.2*
Teacher competence score (% of criteria fulfilled)	66.6	58.3	61.7	-4.9	+3.3
Teacher competence standard fulfilled (%)	67.4	54.8	61.6	-5.8	+6.7
Teacher competence score (% of criteria fulfilled, strict version)		46.0	46.3	n/a	+0.2
Teacher competence standard fulfilled (strict) (%)		9.0	5.2	n/a	-3.8

Note. * indicates statistical significance ($p < .05$). 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%.

We also examine how teachers who reported having received ESSPIN training by CS3 performed compared to those that did not report having received ESSPIN training.

Across most indicators, ESSPIN-trained teachers perform slightly better than non-ESSPIN-trained teachers. Encouragingly, 70% of ESSPIN-trained teachers are using teaching aids interactively (not just reading from or writing on the blackboard), compared to only 46% of non-ESSPIN-trained teachers. In addition, a significantly higher proportion of teachers who have received training are assigning at least two individual or group tasks. ESSPIN-trained teachers also spend less time chanting and somewhat more time explaining; spend more time speaking English; and are more likely to give their learners homework. Although we cannot attribute these results to ESSPIN's intervention, ESSPIN-trained teachers appear to have a better grasp of teaching methods that enhance learning.

On our original measure of teacher competence, ESSPIN-trained teachers appear to be performing worse than non-ESSPIN-trained teachers, but this is influenced to some extent by the measurement of the standard. The indicator on knowledge of the English/mathematics curriculum is only factored into the measurement of the teacher competence standard for teachers who teach either English or mathematics. ESSPIN-trained teachers are more likely to teach English or mathematics than non-ESSPIN-trained teachers. Across the board, all teachers perform poorly on knowledge of the curriculum, but this indicator is factored into the measurement of teacher competence more often for ESSPIN-trained than for non-ESSPIN-trained teachers. In light of this,

we have calculated a teacher competence standard that excludes the indicator of curriculum knowledge. On this standard, a higher proportion of ESSPIN-trained teachers (88%) meet the standard compared to non-ESSPIN-trained teachers (79%), which is what we would expect based on the performance on the individual indicators.

On the strict measure of teacher competence, ESSPIN-trained and non-ESSPIN-trained teachers perform similarly, and very few teachers in either group are able to meet the standard. This again is driven by the poor performance in the content knowledge tests, where the proportion of teachers passing both tests did not differ by ESSPIN training.

Table 22: Kano: Teacher competence in CS3, by intervention group

Intervention group	Non-ESSPIN-trained	ESSPIN-trained	Difference
(1) Knowledge of Eng./mathematics curriculum (%)	31.3	28.1	-3.1
(2a) Use of one or more teaching aid (%)	98.3	99.3	+1.0
(2b) Use of one or more teaching aid, excl. blackboard (%)	45.9	69.5	+23.6*
(3) Praise more than reprimand (%)	79.2	85.4	+6.2
(4) Assigns two or more ind./group tasks (%)	12.8	27.0	+14.2*
(5) Passes English and mathematics test	9.8	11.8	+1.9
Teacher competence score (% of criteria fulfilled)	61.3	62.1	+0.8
Teacher competence standard fulfilled (%)	68.5	53.0	-15.5*
Teacher competence standard (excl. curriculum knowledge, %)	79.2	87.7	+8.5*
Teacher competence score (% of criteria fulfilled, strict version)	42.4	51.1	+8.7*
Teacher competence standard fulfilled (strict) (%)	3.7	7.0	+3.2
Additional indicators:			
Proportion of time spent -- explaining (%)	39.7	45.1	+5.4
-- instructing / presenting / dictating (%)	16.0	18.8	+2.8
-- chanting (%)	10.1	4.8	-5.2*
-- closed question / response (%)	7.2	7.1	-0.1
-- open question / response (%)	3.3	3.0	-0.3
Proportion of time spent speaking English (%)	8.9	17.8	+8.9*
Teacher summarised the lesson (%)	56.4	51.6	-4.7
Teacher revisited the lesson's objectives (%)	18.5	18.6	+0.1
Teacher gave learners homework (%)	14.5	22.3	+7.8*
Teacher tested learners' knowledge (%)	35.5	32.9	-2.5
Teacher marked learners' written work (%)	14.6	15.1	+0.5
Note. * indicates statistical significance (p < .05)			

4.2 Findings from the teacher content knowledge tests

The findings above suggest that teachers' content knowledge worsened between 2014 and 2016. Percentage scores in the teacher content knowledge tests provide a rough indication of teachers' test performance, but analysis using IRT provides more reliable learning scales that can also be interpreted more readily in terms of learning benchmarks (see Allen, 2016a). The teachers' results can be divided into four performance bands in literacy and five performance bands in numeracy. Review of the items that teachers in each band can mostly answer correctly then provides

descriptors for each band (Table 23). For example, a teacher in Band 2 for literacy is one who shows knowledge of some basic phonics, can write a simple sentence, and can perform basic comprehension of a passage, as well as satisfying the easier items – testing limited comprehension of simple passages, basic nouns and verbs – associated with a teacher in Band 1. The teacher in Band 2 cannot typically correctly answer the harder items associated with Bands 3 or 4, such as identifying simple antonyms.

Table 23: Band descriptors based on IRT analysis

Band	Literacy	Numeracy
5		Understands conversion of fractions to decimals, and place values in decimals
4	Creates several sentences, shows knowledge of phonics, punctuation, formal letter layout, suffixes and alphabetical order	Understands ideas of area, nets, pictograms and rounding
3	Past/present of verbs, completes a sentence, extracts basic information from a passage, identifies simple antonyms, forms plurals	Understands basic sets, use of the number line to represent sums, conversion of units of time and mass, can complete word problems involving division
2	Shows knowledge of some basic phonics, writes a simple sentence, basic comprehension of a passage	Simple division, word problems involving addition, signs for arithmetic operations, integer comparisons and integer place values
1	Limited comprehension of simple passages, basic nouns and verbs	Simple addition with a carry, simple subtraction, identifying a fraction, counting, simple regular shapes

Within the literacy and numeracy tests, items can be grouped according to specific sub-domains of learning: reading, writing and grammar within literacy, and number concepts and calculation within numeracy.

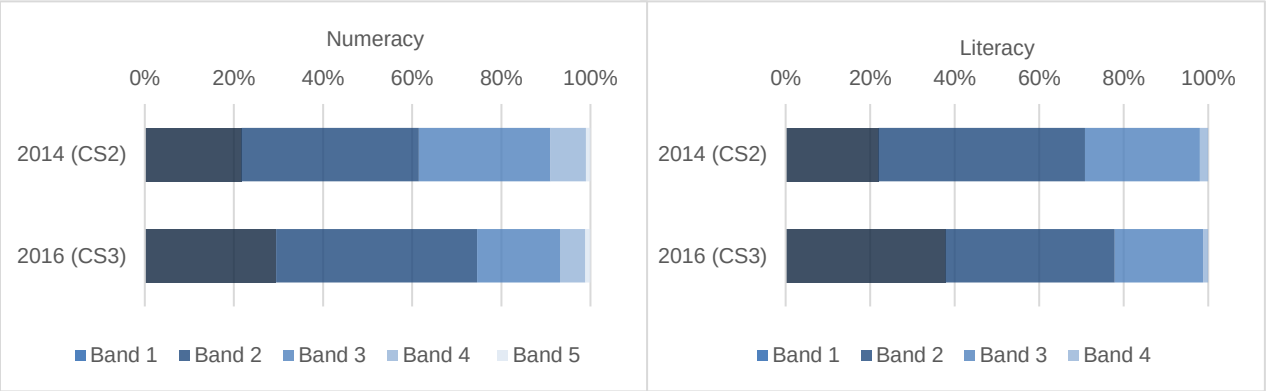
Teachers' scale scores in both English and mathematics have declined significantly since 2014, by about 0.4 of a standard deviation in English and 0.3 of a standard deviation in mathematics (

Table 24). Scores declined across all sub-domains, but declines seemed to be largest for the reading domain in the English tests and for the calculation domain in the numeracy tests. The proportion of teachers in the lowest band increased from 22 to 38% in English, and from 22 to 29% in mathematics; these increases were matched by declines in most of the higher performance bands. Therefore, approximately a third of the teachers in the sample showed only a very limited understanding of literacy and numeracy concepts. The decline in teacher test scores was evident also when only looking at the teacher panel (i.e. those teachers who were tested both in CS2 and CS3), meaning that test performance within the same group of teachers deteriorated over time. It is possible that the change in test administration between CS2 and CS3 may have played a part in the worsening in test performance over time, for example if teachers felt less motivated to complete the tests, which were administered after lessons had finished. Teachers were, however, aware that they would be compensated for their participation in the tests.

Table 24: Kano: Teachers' test scores (IRT analysis) in 2014 and 2016

	2014 (CS2)	2016 (CS3)	Change on average, 2014–16
English IRT scale score (mean 500, s.d. 100)	444	405	**-39.5***
English Band 1 (%)	22	38	+15.9*
English Band 2 (%)	49	40	-8.8
English Band 3 (%)	27	21	-6.3
English Band 4 (%)	2	1	-0.8
Reading (English sub-scale, mean 500, s.d. 100)	450	409	**-40.3***
Writing (English sub-scale, mean 500, s.d. 100)	457	428	**-29.8***
Grammar (maths sub-scale, mean 500, s.d. 100)	442	411	**-31.1***
Mathematics IRT scale score (mean 500, s.d. 100)	441	410	**-31.2***
Mathematics Band 1 (%)	22	29	+7.8
Mathematics Band 2 (%)	40	45	+5.3
Mathematics Band 3 (%)	30	19	-10.9*
Mathematics Band 4 (%)	8	6	-2.3
Mathematics Band 5 (%)	1	1	+0.1
Number concepts (maths sub-scale, mean 500, s.d. 100)	439	415	**-23.4***
Calculation (maths sub-scale, mean 500, s.d. 100)	452	413	**-38.9***
Note. * indicates statistical significance (p < .05)			

Figure 2: Proportion of teachers in each performance band, by year



Teachers who had received ESSPIN training performed better in the content knowledge tests in 2016 than teachers who had not received ESSPIN training. ESSPIN-trained teachers performed about 0.4 standard deviations better on both tests, and also performed significantly better on all sub-scales of the tests. They were more likely to perform in the medium performance bands (Band 3 for the literacy tests and Band 2 for the numeracy tests). Irrespective of ESSPIN training, very few teachers are able to score in the highest performance bands.

Table 25: Kano: Teachers' test scores (IRT analysis) by ESSPIN training

	Non-ESSPIN-trained	ESSPIN-trained	Difference in means
English IRT scale score (mean 500, s.d. 100)	388.2	425.7	**+37.5***
English Band 1 (%)	40.0	35.2	-4.8
English Band 2 (%)	42.6	36.7	-5.9
English Band 3 (%)	16.3	26.7	+10.4
English Band 4 (%)	1.1	1.4	+0.2
Reading (English sub-scale, mean 500, s.d. 100)	389.0	434.9	**+45.8***
Writing (English sub-scale, mean 500, s.d. 100)	418.3	439.3	**+21.1***
Grammar (maths sub-scale, mean 500, s.d. 100)	402.4	421.9	**+19.5***
Mathematics IRT scale score (mean 500, s.d. 100)	389.3	434.0	**+44.7***
Mathematics Band 1 (%)	39.1	18.2	-20.9*
Mathematics Band 2 (%)	39.7	51.4	+11.7*
Mathematics Band 3 (%)	16.3	21.4	+5.1
Mathematics Band 4 (%)	4.4	7.2	+2.8
Mathematics Band 5 (%)	0.6	1.8	+1.2
Number concepts (maths sub-scale, mean 500, s.d. 100)	393.1	441.0	**+47.9***
Calculation (maths sub-scale, mean 500, s.d. 100)	398.7	430.3	**+31.5***
Note. * indicates statistical significance (p < .05)			

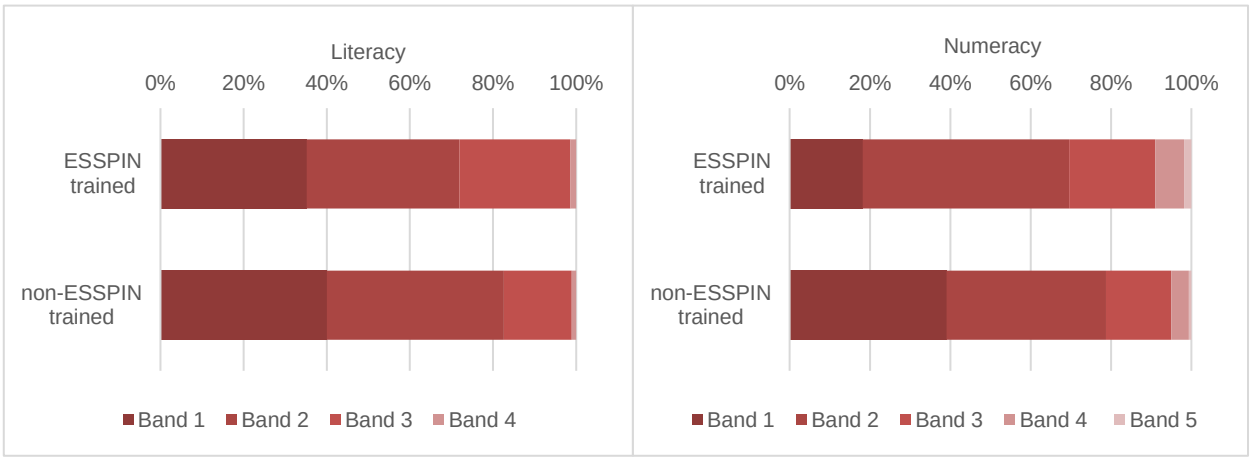


Figure 3: Proportion of teachers in each performance band, by ESSPIN training

4.3 Teacher motivation

Teacher motivation could have a notable influence on the extent to which ESSPIN's contributions to teacher competence translate into improved teaching practices – if teachers are demotivated, they may be less likely to apply these skills or to attend school regularly. Teacher motivation may also be influenced by training and mentoring – as teachers acquire new skills, their self-efficacy may increase, in turn making them more engaged and committed to their jobs.

Box 16: Measuring teacher motivation

For CS3, we included a measure of teacher motivation and teacher interaction using a scale that had been developed for the Nigerian context, and that had been used and tested in two previous school-based surveys. We define teacher motivation as the propensity of teachers to start and maintain behaviours that are directed towards fulfilling their professional goals, and in particular towards achieving better learning outcomes for the school's learners (Cameron, 2015). Many existing instruments designed to measure teacher motivation focus exclusively on 'efficacy' – the extent to which teachers see themselves as able to influence their pupils' learning outcomes – which can also be seen as the 'can do' aspect of motivation (Bennell and Akyeampong, 2007). We wished to go beyond this to include measures relating more closely to teachers' willingness to work hard and their commitment, effort and enjoyment, which might together be labelled as 'will do' aspects of motivation.

The motivation scale we developed was incorporated into the teacher interviews. Teachers were asked to what extent they agreed[9] with a series of statements that measure different aspects of motivation. The scale consists of three sub-scales of teacher motivation (satisfaction, skills and engagement) and one scale of teacher–teacher interaction (collegiality). The three sub-scales of teacher motivation were combined into a composite motivation measure by calculating the mean of the three sub-scales[10]. The teacher motivation scale was also analysed using IRT. Table 26 describes each of the different sub-scales and provides some examples of the items used to assess these.

Source: CS3 Technical Report.

Table 26: Teacher motivation and interaction scale and sub-scales

Scale	Description	Example of items
Collegiality	How I see the extent of commitment and collaboration among my colleagues ('teacher–teacher interaction')	• All of the teachers in my school trust each other • All teachers at this school are highly committed to their job
Satisfaction	The value I place on my role as a teacher ('interest and enjoyment')	• I always enjoy teaching very much • I like to spend a lot of energy to make my classes interesting
Skills	The perception I have of my competences and skills as a teacher ('self-efficacy')	• I believe I know how to teach well • I believe I have the skills needed to encourage my learners to always work hard
Engagement	How engaged and committed I feel I am to my work as a teacher ('pressure/tension')	• It is difficult to manage learners in my classroom • Teaching is very tiring
Composite measure (mean of satisfaction, skills and engagement)		

Table 27 shows the levels of motivation among teachers as reported during CS3, comparing ESSPIN-trained to non-ESSPIN-trained teachers. There are no significant differences in motivation between ESSPIN-trained and non-ESSPIN-trained teachers on any of the indicators. It is

[9] Teachers were asked to pick from amongst the following options: 'strongly disagree', 'disagree', 'agree', 'strongly agree'.
[10] The three sub-scales were also combined into a composite measure using partially non-compensatory methods. These produced composite measures which were very highly correlated with the simple mean composite.

interesting to note that ESSPIN-trained teachers do not have a higher perception of their skills compared to non-ESSPIN-trained teachers, despite our more 'objective' indicators suggesting that ESSPIN-trained teachers are more competent and have better content knowledge.

Table 27: Kano: Teacher motivation and interaction by ESSPIN training

	Non-ESSPIN-trained	ESSPIN-trained	Difference in means
Collegiality	502.3	491.0	-11.3
Satisfaction	467.9	468.7	+0.8
Skills	483.7	475.8	-7.9
Engagement	438.9	452.9	+14.0
Composite motivation measure	465.9	467.4	+1.5

Note. * indicates statistical significance ($p < .05$). All scores are normalised to have an average (mean) of 500 and a standard deviation of 100.

4.4 Summary and discussion

Teacher competence has not been improving over time and performance in the content knowledge tests is worsening significantly over time. The dramatic enrolment increases are likely to have put substantial strain on teachers, who may be less motivated and may find it difficult to apply good teaching practices in overcrowded classrooms. In addition, stakeholders report that the coverage of ESSPIN teacher training has not been sufficient, due to unstable funding, and reductions in the proportion of training to pupil enrolments, which may also provide a possible explanation for why teacher competence has not improved over time.

ESSPIN-trained teachers perform better in the literacy and numeracy content knowledge tests than non-ESSPIN-trained teachers. In addition, ESSPIN-trained teachers are more competent than non-ESSPIN-trained teachers. In particular, we observe ESSPIN-trained teachers performing better in aspects of teaching that require a better understanding of how to enhance learning: for example using teaching aids more interactively and assigning more individual or group tasks during their lessons. While there may be alternative explanations for these findings, this set of results suggests that ESSPIN training in Kano with a particular focus on teaching quality may be having an effect on teachers' content knowledge and competence.

Despite this, teacher performance in the content knowledge tests remains very poor. Although ESSPIN-trained teachers perform better in the tests, they are not more likely to pass both tests. Only 11% of teachers are able to pass both tests in 2016. In addition, the propensity of teachers to assign individual and group tasks is low and there are persisting gaps in the curriculum knowledge of many English and mathematics teachers. The poor performance on these measures of teacher competence is likely to reflect the large gaps in teachers' content knowledge. If teachers have such poor content knowledge, they are likely to have a poor understanding of the curriculum, and to not have sufficient skills to effectively assign individual and group tasks.

5 Trends in school quality

Box 17: School quality: Key findings

- School quality in Kano has been improving over time. In 2016, 17% of schools meet our standard for a good quality school, based on having competent teachers, an effective head teacher, effective school development planning and a functional SBMC. This is a large improvement compared to 2012, when only 3% of schools were of good quality.
- This translates into an estimated 794 more schools meeting the standard for a good quality school in Kano in 2016 compared to 2012. This implies that approximately 431,660 children are learning in a better-quality school.
- However, almost no schools in Kano meet the stricter version of the school quality standard, which also requires half of the teachers to pass a literacy and numeracy test.
- Schools that have received more years of ESSPIN intervention score higher on measures of school quality than schools that have received fewer years of ESSPIN intervention.
- Schools that have received more years of ESSPIN intervention have improved somewhat faster over time, but since this difference is not statistically significant, we cannot with confidence attribute differences in school quality to ESSPIN's intervention.

The ESSPIN logframe defines an overall measure of school quality that draws on the standards for teacher competence, head teacher effectiveness, school development planning and SBMC functionality. A quality school is defined as one that meets the teacher competence standard and at least two of the other standards (Box 18). We also use a 'quality score' indicator, which is an average of the scores that schools achieve on each of the four indicators mentioned above.

Box 18: 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 half the teachers sampled in each school must be competent);
2) head teacher effectiveness standard;
3) school development planning effectiveness standard; and
4) SBMC functionality standard.

The version of this standard used in CS1 did not rely on teacher content knowledge tests. For CS2, we introduce a second, stricter version of the standard, in which teachers must get above 50% in literacy and numeracy tests to be classed as competent.

School quality has increased significantly between 2014 and 2016, with a rise of about 14 percentage points in the school quality score (Table 28). The proportion of schools meeting the overall quality standard as used in CS1 has increased substantially, from around 3% in 2012 to 17% in 2016. This translates into an estimated 794 more schools meeting the standard for a good quality school in Kano in 2016 compared to 2012. This implies that approximately 431,660 children are learning in a better quality school[11].

Almost no schools met the stricter version of the standard in 2014 or 2016. Meeting the stricter standard requires 50% of teachers in each school to have passed the literacy and numeracy test: with only 11% of teachers across all schools meeting this score, most schools fail to meet the

[11] The calculation accounts for the change in the proportion of schools that meet the school quality indicator within each intervention group, and is based on the number of schools and average primary school size by intervention group in Jigawa reported in the Annual School Census 2014/15 (see **Error! Reference source not found.** and **Error! Reference source not found.**).

stricter quality standard purely on the basis of the teacher tests. The stricter school quality standard therefore clearly sets the bar too high to be a useful distinguisher of school quality in Kano schools. However, it does remain a valuable indicator and points to the fact that the quality of learning is likely to be restricted when teachers lack a good grasp of primary-level content.

Table 28: Kano: School quality in 2012–2016

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
Quality score (%)	38.9	34.1	49.1	**+10.2***	**+15.0***
School meets quality standard (%)	3.1	3.3	17.4	**+14.3***	**+14.1***
Quality score (strict version) (%)		31.5	45.4	n/a	**+13.9***
School meets quality standard (CS2 version) (%)		0.2	0.2	n/a	+0.0
Note. * indicates statistical significance (p < .05)					

Schools that have received more years of ESSPIN intervention score significantly higher on the school quality score (both versions) than schools that have received fewer years of ESSPIN intervention (Table 29). The estimated effect of one year of full intervention is an 11 percentage point increase in the chance of meeting the quality standard (CS1 version), but this increase is not statistically significant. There is no effect of the ESSPIN intervention on the proportion of schools meeting the strict quality standard.

Table 29: Kano: School quality by ESSPIN intervention group in 2016

	Min.	Med.	Estimated effect of one year of full intervention
Quality score (%)	47.8	61	**+7.4***
School meets quality standard (%)	14.2	45.6	+10.7
Quality score (strict version) (%)	44	57.8	**+7.8***
School meets quality standard (CS2 version) (%)	0.2	0	n/a
Note. * indicates statistical significance (p < .05)			

Schools in the medium intervention group are scoring higher on the school quality criteria, but this may reflect differences at baseline rather than the contribution of ESSPIN's intervention. To explore this, we analyse changes over time between the minimum and medium intervention groups.

Medium intervention schools had higher quality scores at baseline (CS1), but have also achieved a larger increase in school quality over time: they improved by 14.8 percentage points, compared to minimum intervention schools, which improved by 9.6 percentage points over the period. However, this result is not statistically significant, and we therefore cannot with confidence attribute differences in school quality to ESSPIN's intervention.

Table 30: Kano: Difference in school quality between intervention group, change over time (2012–16)

	Intervention during 2011/12–2014/15		
	0–1 years	**2–3 years**	**Difference**
2012 (CS1)	38.7	48.8	+10.2
2014 (CS2)	33.7	46.4	+12.7
2016 (CS3)	48.3	63.7	+15.4
Difference (2012-2016)	9.6	14.8	+5.2
Note. * indicates statistical significance (p < .05)			

6 Learning outcomes

The ultimate aim of ESSPIN is to improve learning outcomes in government schools in the six states. In this chapter, we examine the trends in learning outcomes over time, and differences in learning outcomes between schools that have received more or fewer years of ESSPIN intervention, and we evaluate whether effects on learning achievement can be attributed to ESSPIN.

Box 19: Learning outcomes: Key findings

- Learner performance in the Grade 4 tests has not changed significantly over time, but learner performance in the Grade 2 tests worsened significantly between 2012 and 2016.
- Learners in schools that have received more years of ESSPIN intervention perform marginally better in the tests than learners in schools that have received little ESSPIN intervention, although results are statistically non-significant on most tests. In the Grade 2 literacy tests, ESSPIN's intervention is associated with significantly better outcomes.
- After controlling for school-level characteristics and past test scores, we do not find evidence to suggest that ESSPIN's intervention has made a positive contribution to learning outcomes in Kano.
- These results have been observed during a period in which total pupil enrolments in Kano have risen by approximately 700,000 pupils. So, whilst the average learning outcomes have stood still or declined, the total 'volume' of education delivered has increased substantially. It is likely that newly enrolled children come disproportionately from disadvantaged or conflict-affected backgrounds, implying that the equity in the provision of education is also likely to have increased over time.

6.1 Pupil learning achievements in English literacy and numeracy

Learning outcomes were measured in literacy and numeracy at Grades 2 and 4, and analysed using IRT (see Allen, 2016b and Allen, 2016c). The analysis for each test produces a scale score, which, by design, has an average (mean) of 500 and a standard deviation of 100. This scale is also divided into bands, indicating the level of proficiency of the learner. For the Composite Surveys, bands have been designed to correspond to the levels of proficiency expected at each grade in the Nigerian curriculum. For example, a learner in Band 2 for literacy is one who is able to demonstrate knowledge and skills in at least some of the tasks that are considered to be within the range of Grade 2 proficiency. Table 31 and Table 32 list some examples of the tasks within each band.

Table 31: Examples of knowledge and skills that learners in each literacy band can demonstrate

Band 4: Grade 4 and above	Read and understand the grammatical structure of a sentence and complete a missing word using 'where', 'which', 'what' and 'who' Follow the conventions of letter-writing to complete a letter template. Completing grammatically accurate sentences, with correct spelling, and a greeting and sign off Read for meaning a short, simple text with a range of sentence structures independently
Band 3: Grade 3 literacy	Read phonically decodable two-syllable and three-syllable words that include common diagraphs and adjacent consonants Independently plan and write a grammatically correct simple sentence Read a simple sentence for meaning and complete a missing word using correct spelling
Band 2: Grade 2 literacy	Use phonic knowledge to utter initial sounds of names of familiar animals Use knowledge of common inflections in spellings, plurals, to write the answer to a question Spell simple high frequency words accurately
Band 1: Emerging literacy	Verbally compose a short grammatically correct sentence in the continuous present tense in response to a question about a picture Listen to a short passage and remember specific details to respond verbally to a question Clearly shaped and correctly orientated copying of words, with an understanding of space and full stops
Band 0: Pre-literacy	Understand and respond verbally with a grammatically correct sentence to a simple question about their age Understand and respond verbally with a grammatically correct sentence to a simple question about their name Use phonic knowledge to utter initial sounds of the names of familiar objects and animals

Table 32: Examples of knowledge and skills that learners in each numeracy band can demonstrate

Band 5: Grade 5 and above	Solve a word problem involving differences in time Determine which number rule was used to make one number into another Solve a simple algebra problem
Band 4: Grade 4 numeracy	Being able to gather information by interpreting simple graphs Calculate the area of a rectangle, multiplying a decimal number, to one decimal place, by a one-digit number, and record the answer in m2 Choose the most appropriate strategy to subtract a decimal number, to two decimal places and a two-digit number, involving measure
Band 3: Grade 3 numeracy	Multiply a two-digit number by a one-digit number Use short division; subtract a two-digit number from a two-digit number crossing the tens boundary Choose a strategy to add a three-digit number and a two-digit number crossing the tens boundary, involving money
Band 2: Grade 2 numeracy	Use non-standard units of measure to compare the capacity of three containers Subtract a two-digit number from a two-digit number Name common 2D shapes Extend counting past 800 and count in tens
Band 1: Emerging numeracy	Recognise and complete a sequence of three two-digit numbers that are multiples of five Subtract a one-digit number from a two-digit number 1–19 Read analogue clock to the hour
Band 0: Pre-numeracy	Compare the length of two straight lines Use non-standard units of measure to compare the capacity of three containers Count to 10

Grade 4 average learning outcomes in Kano have worsened marginally but not significantly over time, and the distribution of students in each performance band has remained similar between 2012 and 2016. In Grade 2, average learning outcomes have worsened over time in both literacy and numeracy tests (Table 33). In the Grade 2 literacy test, a larger proportion of learners are found to be in the lowest performance band in 2016. In the Grade 2 numeracy test, a smaller proportion of learners are in the highest performance band (Band 2) in 2016, but the proportion of learners in the lowest performance band has also decreased, with a consequent increase in those found in Band 1.

In general, learners' test performance remains very poor. In the literacy tests, only between 2% (Grade 2) and 5% (Grade 4) of learners are able to perform at their expected grade level. In the numeracy tests, these proportions are slightly higher, with 6% (Grade 2) and 12% (Grade 4) of learners performing at the expected level or better. However, these results have been observed during a period in which total pupil enrolments in Kano have risen by approximately 700,000 pupils. So, whilst the average learning outcomes have stood still or declined, the total 'volume' of education delivered has increased substantially. It is likely that newly enrolled children come disproportionately from disadvantaged or conflict-affected backgrounds, implying that the equity in the provision of education is also likely to have increased over time.

Table 33: Kano: Learning outcomes in 2012–16

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
Grade 2 literacy score	435.9	455.1	423.8	-12.1	**-31.3***
– Band 0: Pre-school (%)	71.4	74.5	90.2	+18.8*	+15.7*
– Band 1: Grade 1 (%)	10.3	15.5	7.7	-2.5	-7.8*
– Band 2: Grade 2 (%)	18.3	10	2.1	-16.3*	-7.9*
Grade 4 literacy score	438.6	436.7	426.7	-11.9	-10.1
– Band 1: Grade 1 (%)	66.7	77.1	77.9	+11.1	+0.7
– Band 2: Grade 2 (%)	10.2	10.5	12.2	+2.0	+1.8
– Band 3: Grade 3 (%)	4.9	4.5	4.6	-0.2	+0.1
– Band 4: Grade 4 (%)	18.2	7.9	5.3	-12.9	-2.6
Grade 2 numeracy score	487	444.8	423	**-64.0***	**-21.8***
– Band 0: Pre-school (%)	14.8	17.5	9.5	-5.2	-7.9*
– Band 1: Grade 1 (%)	59.1	68.8	84.3	+25.2*	+15.5*
– Band 2: Grade 2 (%)	26.1	13.8	6.2	-20.0*	-7.6*
Grade 4 numeracy score	448.1	442.9	440.6	-7.4	-2.2
– Band 1: Grade 1 (%)	28.7	36.4	37.5	+8.8	+1.1
– Band 2: Grade 2 (%)	43.3	35.7	33.6	-9.7	-2.1
– Band 3: Grade 3 (%)	12	16.5	17.5	+5.5	+1.0
– Band 4: Grade 4 (%)	9.4	8.2	9	-0.4	+0.8
– Band 5: Grade 5 (%)	6.7	3.3	2.5	-4.2	-0.7
Note. * indicates statistical significance (p < .05)					

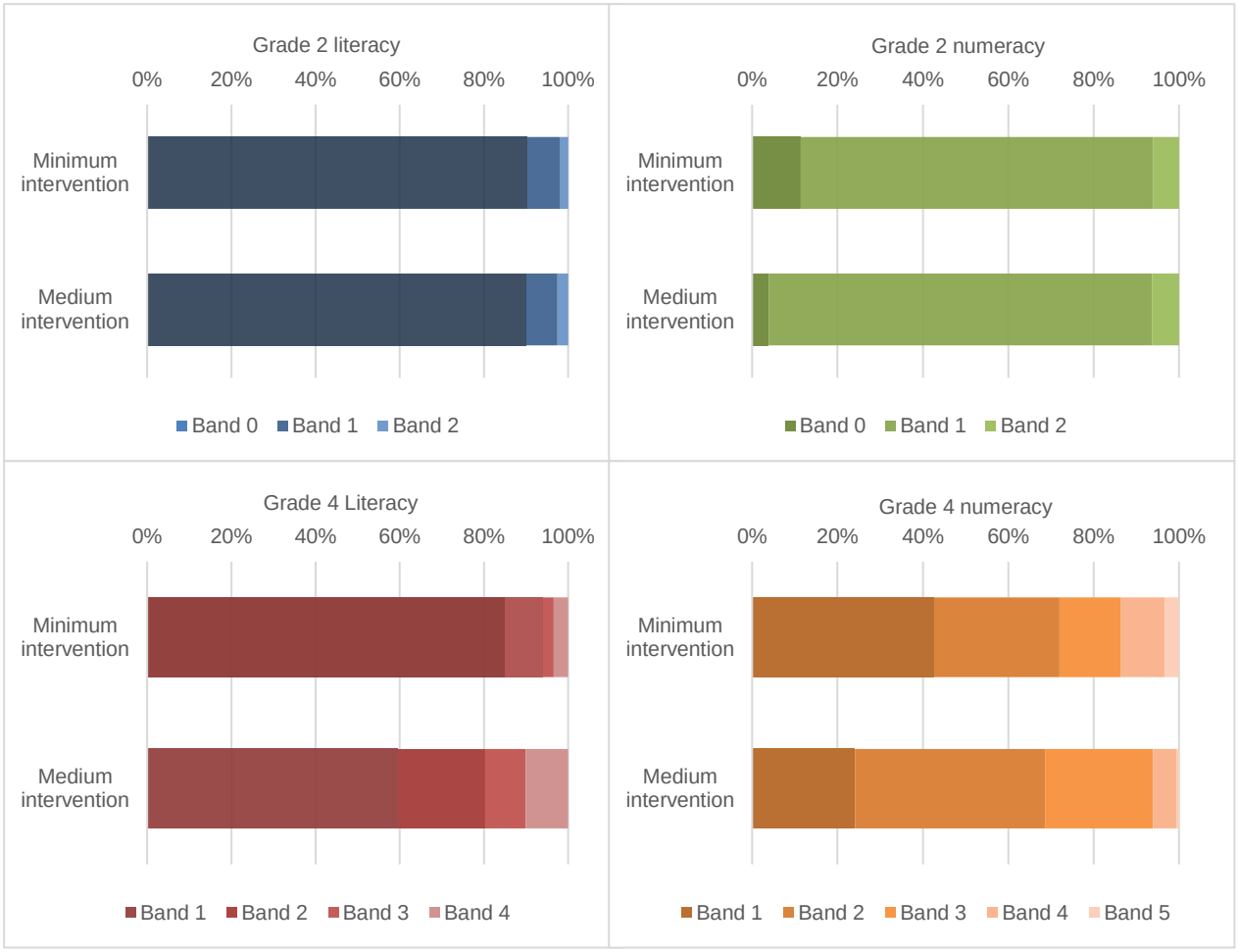
Looking at the different intervention groups, learners in schools that have received more years of ESSPIN intervention appear to be performing slightly better than learners in schools that have received fewer years of ESSPIN intervention, although the estimated effect of one year of full ESSPIN intervention is not significant for three of the tests. In the Grade 2 literacy tests, for each additional year of intervention the school has received, learners' performance is estimated to increase by about 0.14 standard deviations, which is statistically significant.

Learners in schools which have received minimum intervention from ESSPIN are disproportionately concentrated in the lowest performance bands, while those in schools which have received medium intervention are more likely to be found in the medium bands (Figure 4). Very few learners reach the highest achievement bands in either intervention category.

Table 34: Kano: Learning outcomes by ESSPIN intervention group in 2016

	Min.	Med.	Estimated effect of one year of full intervention
Grade 2 literacy score	419.3	438	**+13.6***
Band 0: Pre-school (%)	90.2	90	-2.2
Band 1: Grade 1 (%)	7.8	7.5	+1.2
Band 2: Grade 2 (%)	1.9	2.5	+0.9
Grade 4 literacy score	418.8	447.2	+14.4
Band 1: Grade 1 (%)	84.9	59.6	-10.4*
Band 2: Grade 2 (%)	9	20.6	+6.3*
Band 3: Grade 3 (%)	2.7	9.7	+1.9
Band 4: Grade 4 (%)	3.4	10.1	+1.8
Grade 2 numeracy score	419.8	432.9	+11.8
Band 0: Pre-school (%)	11.3	3.9	-7.2
Band 1: Grade 1 (%)	82.5	89.8	+4.3
Band 2: Grade 2 (%)	6.1	6.3	+0.9
Grade 4 numeracy score	436	452.7	+7.5
Band 1: Grade 1 (%)	42.6	24.1	-8.6
Band 2: Grade 2 (%)	29.4	44.5	+6.8
Band 3: Grade 3 (%)	14.4	25.3	+3.7
Band 4: Grade 4 (%)	10.3	5.6	-1.7
Band 5: Grade 5 (%)	3.3	0.5	-2.9
Note. * indicates statistical significance ($p < .05$)			

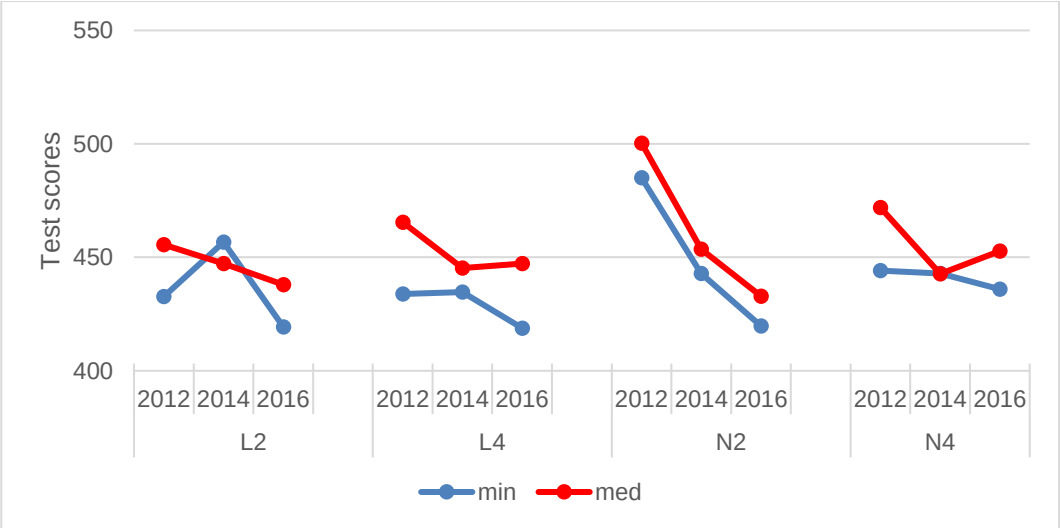
Figure 4: Kano: Distribution of test scores by intervention group in 2016



In Figure 5, we disaggregate the change over time according to ESSPIN intervention. Schools that received more years of ESSPIN intervention had better learning outcomes at baseline in 2012 than schools that received fewer years of ESSPIN intervention. The overall trend seems to suggest that average learning outcomes are worsening slightly within a rapidly increasing enrolment, at about the same rates in the minimum and medium intervention groups—at least when comparing scores in 2012 and scores in 2016 (there are different trajectories in 2014 across the different tests).

These patterns are explored more rigorously in Section 6.2 below, using regression and matching analysis to examine how change over time varies with ESSPIN intervention, and controlling for possible confounding variables, such as school characteristics.

Figure 5: Learning outcomes by test, year and ESSPIN intervention group



6.2 Controlling for school and pupil characteristics

6.2.1 Differences in background characteristics

As has been noted in Section 1.3 and Annex A, schools in the minimum and medium intervention groups in Kano differ from each other in regard to several background characteristics. The schools that have received more years of ESSPIN intervention have been established for longer, have better infrastructure in their classrooms, and have more qualified teachers and lower PTRs. They are also much larger and more likely to be located in urban areas. Urban schools, in particular, tend to have better learning outcomes than rural schools, and if left uncorrected, this difference between the intervention groups could bias our estimates of ESSPIN intervention effects upwards.

Schools that have received more years of ESSPIN intervention also show significantly larger increases in enrolment and PTRs. Schools may have difficulty coping with rapid enrolment increases, and increases in the PTR would tend to reduce teachers' ability to ensure all students achieve good learning outcomes. Therefore, this difference between intervention groups in Kano could bias our estimates of ESSPIN intervention effects downwards.

There are a number of differences between the groups of schools that have had more years of ESSPIN intervention and those that have had fewer years of intervention, and taken together these could bias our estimates of ESSPIN's effect in either direction. We use a number of statistical methods to control for these differences in the following sections.

6.2.2 Timing of ESSPIN intervention and learning outcomes in 2016

Analysing the effects of ESSPIN's intervention is made complicated by the variations in the timing and duration of the intervention for different schools in Kano. The initial pilot schools in Kano received two years of 'full package' intervention, but then received no further intervention for two years, before receiving the intervention again under the scaled-up model. A further group of schools has consistently received the intervention since 2012/2013, while the largest group of schools has consistently received the intervention since 2013/2014.

The timing of the intervention could affect learning outcomes in at least two ways. First, it may take time for the effects of leadership training, teacher training and other interventions to filter through to measurable gains in learning outcomes. This is especially so given that children are tested in the

second and fourth grades. From this perspective, for example, a Grade 4 learner whose teachers benefited from training during the current school year is unlikely to benefit as much as one whose teachers were trained four years ago, so that they have had consistent exposure to better teaching. This would suggest that training initiated a longer time ago would have larger effects on learning outcomes. Second, however, it is possible that effects of training wear off over time, particularly if it is common for teachers to transfer between schools or if trained teachers retire. More recent training would then have a larger effect on learning outcomes.

We explore this using a regression analysis that compares every combination of timing and duration of intervention (with the average across these intervention patterns serving as a baseline for comparison). In addition to a simple regression model, we report a model that also controls for differences in school characteristics (Table 35).

Learning outcomes across the different timings of the intervention are very similar. The initial pilot schools have somewhat higher learning outcomes, but these differences disappear once we control for the background characteristics of the schools. This result is perhaps unsurprising since in the CS2 report we also found no significant differences between the intervention groups, and since then all schools have received the same amount of intervention. The earlier years of intervention in the initial pilot schools may either have been insufficient to influence learning outcomes or the effect may have worn off over the two year hiatus of intervention.

Table 35: Kano: Difference in test scores in 2016 by timing of ESSPIN intervention

Years of intervention	Total years	L2		L4		N2		N4	
2013/14 and 2014/15 only	2	-4.9		-4.0		-4.6		-6.3	
2012/13, 2013/14, 2014/15	3	-4.9		-6.7		-12.0		+1.2	
2009/10–2010/11 and 2013/14, 2014/15	4	+9.8		+10.7		+16.6		+5.2	
Controlling for background characteristics									
2013/14 and 2014/15 only	2	+4.7		+7.5		+6.4		+1.2	
2012/13, 2013/14, 2014/15	3	-1.3		-7.9		-14.9		+4.6	
2009/10–2010/11 and 2-13/14, 2014/15	4	-3.3		+0.4		+8.5		-5.8	
Note. * indicates statistical significance (p < .05)									

6.2.3 Are learning outcomes better in schools that have received more years of intervention in 2016?

In Section 6.1 above we found that learners from schools that had received more years of intervention had better learning outcomes than those from schools that had received fewer years of intervention, although these results were only statistically significant for the Grade 2 literacy test. However, there are also some pre-existing characteristics of the schools that received more years of intervention which might have biased these results upwards (see Annex A). Therefore, in this section, we add statistical controls for these 'confounding variables' – characteristics of schools that might affect learning outcomes and make it harder to tell whether the intervention is having an effect or not. We also estimate a model which controls for pre-existing differences in test scores by adding test scores in CS1 as a confounding variable.

We use ordinary least squares regression analysis to estimate the models. Regression analysis estimates the correlation of learning outcomes with ESSPIN's intervention, conditional on school characteristics and pre-existing differences in test scores. We present four different models that control for a range of pre-existing differences between schools and learners.

None of our models find a significant effect of ESSPIN's intervention on learning outcomes in Kano (Table 36). The initial significant effect of ESSPIN's intervention on Grade 2 literacy outcomes no longer persists once we control for background characteristics. The results are also no longer all in the expected direction: depending on the model used the effect of ESSPIN's intervention is either slightly positive or slightly negative. This confirms the results shown in Figure 4, which suggested that learning outcomes in both the minimum and medium intervention groups follow similar patterns over time.

Table 36: Kano: Estimates of the effect of ESSPIN intervention on learning outcomes in 2016

Model	L2		L4		N2		N4	
(1) Simple regression, clustered standard error (SE), no sample weights	+9.28		+7.93		+10.68		+9.93	
(2) Full covariates	-7.35		-9.99		-5.28		-3.4	
(3) Lagged school-level learning outcomes	-3.78		+1.67		+4.75		+2.89	
(4) Lagged outcomes and covariates	-4.91		+0.25		-3.13		+1.47	
Note. * indicates statistical significance (p < .05)								

6.3 Summary and discussion

We have applied IRT to measure learners' performance in literacy and numeracy, in Grades 2 and 4. Learning outcomes worsened significantly over time in Grade 2, while they did not change significantly in Grade 4. There are a few potential explanations for this deterioration in learning outcomes. First, poor content knowledge amongst teachers is likely to be a key constraint, and it is notable that this has also worsened between 2014 and 2016. It is questionable how much impact one can hope to see from improved school management if teachers' mastery of Primary Grade-level content is so limited. Secondly, school enrolment and PTRs have increased substantially, and learners are likely to face a more difficult learning environment, for example due to overcrowding of classrooms. In addition, the profile of learners at public schools in the state may be changing. It is possible that enrolment increases in Kano in recent years are partly due to increasing enrolment of previously excluded pupils, who are likely to be from more disadvantaged and displaced backgrounds, and to be less well prepared to enter the primary school system.

Learning outcomes appear to be better for learners whose schools have received more years of ESSPIN intervention. For all four tests, the estimated effect of spending time in a school which has had two or more years of ESSPIN intervention is positive, but it is only statistically significant for the Grade 2 literacy test.

There are some significant pre-existing differences in the schools that have received more years of ESSPIN intervention and those that received fewer years of intervention. For example, schools that received more years of intervention are more likely to be urban. We used a number of statistical methods to control for these differences and reduce bias in our estimates of the effect of ESSPIN's intervention.

After controlling for background characteristics, we find no effect of ESSPIN's intervention on learning outcomes, either by the timing of the intervention or by intervention group. It is important to remember that 34 of the schools in our medium intervention group sample (the initial pilot schools) have received the same level of ESSPIN intervention for the past five years as schools in the minimum intervention group. Only 20 schools in our medium intervention group sample received some additional intervention in 2012/2013. At the same time, schools in the medium

intervention category have been faced with larger enrolment and PTR increases than schools in the minimum intervention group. This may explain to some extent why we do not find an effect of ESSPIN's intervention on learning outcomes.

7 Conclusion and implications for ESSPIN in Kano

The Composite Surveys' findings paint a mixed picture with regards to the progress made on school-level outcomes in Kano State. These findings occurred in the context of dramatic enrolment increases, with approximately 700,000 additional learners entering the schooling system between 2009 and 2016. This is likely to have put pressure on head masters and school systems, and to have made teaching conditions more difficult for teachers. It also means that even where learning outcomes stand still or decline slightly, a larger absolute number of children are better educated.

Since 2012, some indicators have improved significantly (notably SBMC functionality and SDP effectiveness), some have remained level (head teacher effectiveness), and others have worsened (school inclusiveness). While head teachers in Kano are not more effective overall, they are more likely to have carried out lesson observations and professional development meetings in 2016. These two indicators seem to have been a particular focus of the leadership training in Kano since 2012.

Across all school management indicators, a high proportion of schools in Kano still do not meet ESSPIN's standards for a good school. In 2016, 17% of schools meet the standard on head teacher effectiveness; 17% on school development planning; 22% on inclusion; 43% on functional SBMCs; and 17% or 0.2% on overall school quality (depending on which indicator we use).

Despite mixed performance over time, schools that have received more years of ESSPIN intervention perform better on most indicators than schools that have received fewer years of intervention. They are of a better school quality, are much better at school development planning, are more inclusive, and are more likely to have well-functioning SBMCs in which women and children participate. Head teachers in schools that have received more years of ESSPIN intervention do not perform better than those in schools that have received fewer years of intervention, but this may in part be explained by the fact that since CS1 most schools had received the same amount of leadership training (nine days in 2013/2014), and that leadership training has not been a strong focus in Kano in recent years.

The findings for teachers are consistent with Kano's decision to shift the focus of ESSPIN's activities in the state to improving teaching quality. ESSPIN-trained teachers perform much better in the literacy and numeracy content knowledge tests than non-ESSPIN-trained teachers. In addition, ESSPIN-trained teachers are more competent and demonstrate a better understanding of how to enhance learning. This set of results suggests that ESSPIN training in Kano with a particular focus on teaching quality may be having an effect on teachers' content knowledge and competence.

Learning outcomes in Grade 2 have worsened significantly since 2012, while learning outcomes in Grade 4 have not changed significantly. ESSPIN's intervention is associated with better learning outcomes, but this is only significant for the Grade 2 literacy test. This significant effect disappears once we add controls for school characteristics and past test scores. Therefore, after taking into account pre-existing differences between schools, we find no effect of ESSPIN's intervention on learning outcomes.

Trends in learning outcomes may be linked to the deterioration in teachers' content knowledge over time. Only 11% of teachers can pass both content knowledge tests, and ESSPIN-trained teachers are no more likely to pass the tests than non-ESSPIN-trained teachers. These findings raise potential questions about the extent to which improvements on other school management indicators can contribute to improved learning in the face of fundamental weaknesses in teachers' own knowledge. This is likely to remain a key barrier to achieving improvements in learning

outcomes, especially during periods of rising learner enrolment and declining investments in the professional development of teachers.

Bibliography

Bennell, P., and Akyeampong, K. (2007) 'Teacher Motivation in Sub-Saharan Africa and South Asia (Department for International Development: Educational Papers)'. DFID.

Cameron, S. (2015) 'Teacher motivation in the TDP baseline survey'. Unpublished technical note. Oxford Policy Management.

Coinco, E. (2014) 'Education, Conflict and Violence Research'. ESSPIN Report 428.

ESSPIN (2013) 'Extension of the Education Sector Support Programme in Nigeria, August 2014 – January 2017'. Business case for DFID.

Megill, D. (2014a) 'Recommendations on Sampling Plans for the Second Composite Survey and the GEP-3 Unified Survey (GUS) in Nigeria'. Unpublished note written for ESSPIN and OPM.

Megill, D. (2014b) 'Final sample design and weighting procedures for ESSPIN Second Composite Survey (CS2) in Nigeria'. Unpublished note written for ESSPIN and OPM.

Mezger, C. (2016) 'Trends of inequality in access to schooling in Nigeria. Evidence from three DHS surveys (2003, 2008, 2013)'. Abuja: Education Data, Research and Evaluation in Nigeria (EDOREN)

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

Sanni, K. (2015) 'Taking School Improvement to Scale: the Education Sector Support Programme in Nigeria'. Cambridge Education.

Annex A School characteristics

The table below sets out summary statistics for Kano's schools, split by categories according to the level of Output Stream 3 intervention (minimum, medium). The data come from the Annual School Censuses from 2009/10, 2013/14 and 2014/15.

Kano's schools by level of ESSPIN intervention	Total	Min.	Med.	
Distance from local government authority headquarters	8.7	8.7	8.7	
Age of the school in 2014	22.9	21.7	33.3	*
Urban (%)	35.8	33.8	53.0	*
Nomadic (%)	3.3	3.6	1.3	*
Islamic (%)	57.1	59.6	35.1	*
Double shift (%)	9.9	8.5	21.5	*
Had parent–teacher association in 2014/15 (%)	97.6	97.8	96.5	*
Had SBMC in 2014/15 (%)	91.9	91.6	95.0	*
PTR in 2009/10	62.9	64.2	52.0	*
PTR in 2013/14	74.6	75.8	63.1	*
PTR in 2014/15	69.3	70.0	63.2	*
% change in PTR between 2009/10 and 2013/14	66.4	63.3	90.6	*
% change in PTR between 2013/14 and 2014/15	11.7	12.3	6.5	
Number of classrooms in 2014/15	5.8	5.3	10.4	*
Number of teachers in 2014/15	8.8	7.8	17.3	*
Primary enrolment in 2009/10	399.2	348.3	833.4	*
Primary enrolment in 2013/14	444.3	382.8	1,036.7	*
Primary enrolment in 2014/15	461.7	396.9	1,024.3	*
% change in enrolment 2009/10–2014/15 (%)	41.8	40.5	51.8	*
% change in enrolment 2013/14–2014/15	8.6	9.5	0.8	*
% of teachers with academic diploma/degree	61.7	61.1	66.3	*
% of teachers with PGDE, BEd or MEd	8.0	7.9	9.0	*
% of teachers with NCE, Grade II or equivalent	54.4	53.6	61.2	*
School has a power source (grid/other)	38.1	37.0	47.3	*
% of classrooms with enough seating	32.1	31.4	34.9	
% of classrooms with a good blackboard	70.3	68.9	76.2	*
% of classrooms in good condition/minor repairs	78.2	77.3	82.0	*
School has at least one toilet (%)	37.6	37.1	41.0	
Number of schools	5,482	4,973	509	

Notes: (1) * indicates a significant coefficient when running a linear or logistic regression of the variable of interest (dependent variable) on the number of years of ESSPIN intervention (independent variable); (2) the 'total' column includes schools that do not have an intervention code. (3) the PTRs shown in the table are calculated as the average PTRs for schools in the state ($\frac{\sum {}^{P_i}/_{T_i}}{N}$) and not the PTR for the state as a whole ($\sum {}^{P_i}/_{\sum T_i}$).

Annex B ESSPIN Output Stream 3 interventions

The table below shows the ESSPIN Output Stream 3 interventions delivered to date in Kano. 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 full intervention received before 2015/2016. Full intervention means the school received some leadership training, some teacher training, and some school visits during the year, though the amount of each may vary. The schools have been grouped as follows: minimum (zero to one years), medium (two to three years), maximum (four to five years).

	Category (years of intervention)	2009/2010			2010/2011			2011/2012			CS1	2012/2013			2013/2014			CS2	2014/2015			2015/2016			CS3
		L	T	SV	L	T	SV	L	T	SV		L	T	SV	L	T	SV		L	T	SV	L	T	SV	
Kano	Minimum (0a)																						20	8	
	Minimum (0b)																			4	6		20	8	
	Minimum (1)														9	9	9			4	6		20	8	
	Medium (2)											6	3	9	9	9	9			4	6		20	8	
	Medium (3)	5*	5*	9*	10*	5*	9*								9	9	9			4	6		20	8	

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 Kano under different headings: SBMC training; women's and children's participation training; and mentoring visits.

	Intervention Category	2010/2011			2011/2012			CS1	2012/2013			2013/2014			CS2	2014/2015			2015/2016			CS3
		S	P	M	S	P	M		S	P	M	S	P	M		S	P	M	S	P	M	
Kano	No intervention															2						
	Post-CS1											7	0	4		1	2	6	1	2	4	
	Post-CS1								1	6	4	0	0	4		3	2	8	1	2	4	

Note: S = SBMC training; P = women's and children's participation training; M = mentoring visits

Annex D Difference-in-difference analysis using regressions

Test	Treatment variable	Model	Coefficient	SE	P value	N	R-squared
L2	pu_exposure	Simple model with survey weights	6.8	2.22	0.003	691	0.042689
L2	pu_exposure	No survey weights but clustered SEs	11.56	1.43	0	2836	0.064225
L2	intervention_binary	Binary exposure variable	9.28	8.15	0.256	691	0.006301
L2	pu_exposure	Full covariates, survey weights	0.24	1.92	0.9	553	0.196005
L2	pu_exposure	Full covariates, no weights	-2.04	2.02	0.314	549	0.196868
L2	intervention_binary	Full covariates	-7.35	6.64	0.27	549	0.196894
L2	pu_dexp13	Lagged school-level learning outcomes	-3.78	5.5	0.494	385	0.161261
L2	pu_dexp13	Lagged outcomes and covariates	-4.91	5.76	0.397	321	0.204887
L4	pu_exposure	Simple model with survey weights	3.23	2.22	0.148	669	0.014037
L4	pu_exposure	No survey weights but clustered SEs	6.78	0.64	0	3202	0.098513
L4	intervention_binary	Binary exposure variable	7.93	8.85	0.371	669	0.00275
L4	pu_exposure	Full covariates, survey weights	-0.07	1.24	0.955	511	0.229306
L4	pu_exposure	Full covariates, no weights	-0.91	1.22	0.456	506	0.285318
L4	intervention_binary	Full covariates	-9.99	8.16	0.223	506	0.287152
L4	pu_dexp13	Lagged school-level learning outcomes	1.67	2.3	0.471	369	0.152293
L4	pu_dexp13	Lagged outcomes and covariates	0.25	2.62	0.925	293	0.291232

N2	pu_exposure	Simple model with survey weights	5.9	3.62	0.105	690	0.008853
N2	pu_exposure	No survey weights but clustered SEs	11.62	1.47	0	2801	0.057978
N2	intervention_binary	Binary exposure variable	10.68	9.96	0.285	690	0.003124
N2	pu_exposure	Full covariates, survey weights	-0.35	3.11	0.911	551	0.109282
N2	pu_exposure	Full covariates, no weights	0	3.03	0.999	547	0.146023
N2	intervention_binary	Full covariates	-5.28	10.98	0.631	547	0.146602
N2	pu_dexp13	Lagged school-level learning outcomes	4.75	9.2	0.607	385	0.078968
N2	pu_dexp13	Lagged outcomes and covariates	-3.13	10.79	0.772	321	0.165954
N4	pu_exposure	Simple model with survey weights	1.56	1.7	0.359	670	0.00241
N4	pu_exposure	No survey weights but clustered SEs	6.9	0.7	0	3177	0.085963
N4	intervention_binary	Binary exposure variable	9.93	8.77	0.259	670	0.003173
N4	pu_exposure	Full covariates, survey weights	-1.76	1.52	0.249	522	0.118532
N4	pu_exposure	Full covariates, no weights	-0.86	1.53	0.576	518	0.146694
N4	intervention_binary	Full covariates	-3.4	10.66	0.75	518	0.146096
N4	pu_dexp13	Lagged school-level learning outcomes	2.89	2.3	0.211	368	0.059754
N4	pu_dexp13	Lagged outcomes and covariates	1.47	3.04	0.631	299	0.132315