Chapter 7
The Impact of Federations on Student Achievement

7.1 Introduction

In the previous chapter we introduced the concept of a federation, a formal collaborative arrangement between schools that statutorily exists in the English system. In that, and previous, chapters, we have also touched quite a number of times on the perceived advantages of networks and collaboratives in education. However, the bottom line for any educational intervention must be the extent to which it has a positive impact on pupil outcomes. Of course, such outcomes can be varied. Self-esteem, well-being, citizenship, and a disposition towards lifelong learning have all, at various times, been mentioned as goals of schooling and education (Muijs & Reynolds, 2010). However, there exist to our knowledge no education systems where the acquisition of cognitive skills as measured through pupil achievement is not a core goal of the system. The question therefore arises of the extent to which networking and collaboration may positively impact on these outcomes. We have discussed what evidence there is on this in Chapter 1, but as we pointed out in that chapter there is a dearth of evidence on impact on pupil learning and achievement. In this chapter we will explore this question by looking at the relationship between being part of a federation and pupil achievement in English schools, using quantitative methodologies.

7.2 Methods

7.2.1 Researching the Relationship Between Federations and Student Outcomes: Aims and Objectives

The aim of this study is to investigate the ways in which federations seek to improve student outcomes and leadership capacity and the extent of variation in their impact and abilities to promote change. The study also explores factors that facilitating positive impacts of federations and any that act as barriers to improvement and examine whether some models are more effective than others in promoting better outcomes. The specific objectives of this study are to:

D. Muijs et al., *Collaboration and Networking in Education*,
DOI 10.1007/978-94-007-0283-7_7, © Springer Science+Business Media B.V. 2011

1. Investigate changes in student outcomes for federation in terms of key attainment indicators and value added measures and compare these to national trends and results for schools with similar intake characteristics;
2. Assess the impact of federations on leadership capacity and effectiveness; and
3. Assess the impact of federations on the quality of teaching and learning.

7.2.2 Researching the Relationship Between Federations and Student Outcomes: Defining a Research Approach

Our broad approach has been to develop a matched sample of schools in federations and non-federations and conduct quantitative analyses to obtain measures of the apparent impact federations have had on the changes in the educational outcomes of different groups of students over the time (1–4 years). Our approach has involved identifying schools that federated in academic years 2007–2008, 2006–2007, 2005–2006, 2004–2005, and 2003–2004 and comparing the examination performance of cohorts prior to federating with examination performance since federation in terms of key indicators such as %5+ A*–C with and without English and Mathematics, and in value added and contextual value added. In addition, each federation has been matched to a comparator 'non-federation' (in terms of key statistics including those related to pupil intake, rural/urban location, previous attainment profiles) to assess the impact of federating. The following section details the methods used to collect and analyse the data.

7.2.3 Methods

A quantitative methodology was used to explore the relationship between school federations and student performance. National pupil and school level datasets were collected from the Department for Children Schools and Families (DCSF) (now department for Education) to allow us to look at performance measures controlled for student background over time. PLASC and National Pupil Database data were requested from and provided by DCSF for this purpose. Data were collected for each year from 2001 to 2008. Ofsted inspection gradings were provided by the National College for School Leadership. As no definitive central list of federations existed when the study was conducted, a random sample of 50 Local Authorities was selected. Each local authority was contacted by the members of the research team with the request to identify federations and the schools that were a part of them. A total of 264 schools and 122 federations were identified in this way.

Follow up telephone calls were made to each of the schools to ascertain whether there were any errors in the designation of the school as a federation and to collect additional data on date the schools federated, the number of heads, and federation structures. A number of schools/federations were identified that had ceased to operate or did not fit the criteria for federation. These were replaced with other schools/federations.

In order to look at the impact of federation on performance, we opted for a quasi-experimental design where each federation school was matched using propensity score matching to a school as similar as possible on key characteristics prior to federating. National datasets were used to match schools on a number of criteria, including:

1. Phase (e.g. primary, middle, and secondary).
2. Type of school (e.g. Voluntary Aided, Voluntary Controlled, and Academy).
3. Gender intake (co-educational, single sex boys, and single sex girls).
4. Performance levels (e.g. percentage achieving KS threshold levels in English and Maths).
5. Pupil intake characteristics (percentage pupils identified as having Special Educational Needs, percentage pupils eligible for Free School Meals).
6. Location (this measure went beyond traditional rural/urban identification, and attempted to match areas that were as similar as possible on socio-demographic characteristics. For example, Cambridge would be matched to York and Salford to Gateshead).
7. School size (as indicated by pupil roll).

Clearly, no schools could be matched identically on these criteria. However, as close a match as possible was sought in all cases. As with the federation schools, all comparator schools were contacted by a member of the research team to ascertain that they were not themselves in a federation and to collect data on Headship and governance. As a result of this, a number of schools had to be replaced as they were themselves part of a federation or had ceased to operate. A range of quantitative methodologies were used to analyse the data (see results section), including univariate and multivariate statistics and multilevel modelling. The Stata and MLWin software packages were used for these analyses.

7.2.4 Sampling

The final sample contained a total of 50 LAs, and 264 schools. These are grouped into 122 federations. Two hundred sixty-four comparator schools were selected to match these. A total of 88.1% of schools in the sample belonged to a two-school federation, 8.5% were part of a three-school federation, with the remainder being part of larger federations. The distribution of schools across phases was (Table 7.1):

A total of 11.3% of schools were Catholic, 16.2% were CofE, and 4.1% were Academies.

The federations tended to have been formed relatively recently, reflecting a rapid development in this area (Table 7.2):

Most federations in the sample were formed in either 2007 or 2008, with just over a quarter formed in 2005 or earlier. Of total federations surveyed, 80.97% had

Table 7.1 Distribution of schools across phases in percentages

Nursery	3.8
Infant	7.2
Junior	7.2
Primary	39.6
First	9.0
Middle	5.3
High	1.9
Secondary	22.3
Special	3.7

Table 7.2 Formation year

Year formed	Percentage
2001 and earlier	1.79
2002	0.00
2003	0.00
2004	1.79
2005	4.46
2006	17.86
2007	32.59
2008	33.48
2009	8.04

a joint head teacher, 19.03% had not. Of total federations surveyed, 14.6% had a joint governing body, 85.6% had not (Table 7.3).

Almost half of the schools surveyed had had only one Head in the past 5 years, with over a third having had two. There were very few differences between federation and comparator schools in this regard.

Table 7.3 Number of head teachers in the past 5 years

Number of heads in last 5 years	% Federation schools	% Comparator schools
1	49.55	49.54
2	35.59	38.99
3	10.36	9.17
4	3.6	1.94
5	0.9	0.46

7.2.5 Federation and Comparator School Characteristics

Federation and comparator schools were compared on key intake variables. Schools were exactly matched on:

– School type
– Gender intake
– Phase

Table 7.4 Characteristics of federation and comparator schools on key variables

	Federation mean	Comparator mean	*t*
School roll	287.9	281.4	−0.74
Percentage pupils with SEN	30.2	31.6	1.35
Percentage reaching threshold targets in English	79.2	*77.9*	0.98
Percentage reaching threshold targets in Maths	74.3	76.8	1.96*

* Significant difference at the 0.05 level.

Table 7.4 depicts match for the variables school roll, SEN percentage, FSM percentage.

The only significant difference found was on Maths achievement, where comparator schools significantly outperformed federation schools, though the difference was small.

7.3 Results

7.3.1 A Typology of Federations

The data was interrogated to find types of federations. Six different types of federations were identified:

a. *Cross-Phase Federations* – Federations consisting of two or more schools of different phases, e.g. a primary and secondary school, or a First, Middle, and High school. Of total schools in the sample, 35.1% are part of this type of federation.
b. *Performance Federations* – Federations consisting of 2 or more schools, some of which are low performing, and others high performing, usually two schools. Of total schools, 15.6% are part of this type of federation.
c. *Size Federations* – Federations consisting of two or more very small or small schools, or a small school and a medium-sized school. Of total schools, 18.8% are part of this type of federation.
d. *Mainstreaming Federations* – One or more special schools combine with one or more mainstream schools. Of total schools in the sample, 4.6% are part of this type of federation.
e. *Faith Federations* – Two or more schools of the same denomination combine. This type can overlap with one of the other four types, but in many cases does not. Of total schools in the sample, 14.8% are part of this type of federation
f. *Academy Federations* – Two or more Academies run by the same sponsor form a federation. Of total schools, 2.3% are a part of this type of federation.

A total of 7.8% of schools were in federations that were not immediately classifiable.

7.3.2 *Federations and Impact on Performance*

Multilevel statistical models were used to look at the impact of federation on performance. This was done only for the cohorts of federations formed in 2005 and earlier and 2006, as no impact is to be expected for those federations formed in 2007 and later in light of previous research on the length of time it takes for federations to become fully operational (see Lindsay et al., 2005). Levels were school (level 2) and pupil (level 1). As the data relates to different cohorts in different years analysis of each year was done separately.

Models were tested for the year of formation and 3 years prior data combined, and for subsequent years up to 2008. A null model was formulated with no predictors. In the next model federation was added, while in the final model for each year other correlates of achievement were included, such as gender, SEN status, and FSM eligibility. Outcome variables were pupil-level achievement, such a KS levels or % 5A*–C grades. As our variable of interest was a school-level variables, all predictors in the analyses are school-level variables.

7.4 2005 Cohort

In the 2005 cohort analyses are only presented for the primary phase (KS2), due to the fact that the sample of secondary schools for this cohort was at four (making a total of eight schools including comparators) too small to provide stable estimates.

Tables 7.5, 7.6, and 7.7 show that for the 2005 cohort there is some evidence of impact of federations over time. Overall, the majority of the variance in both English

Table 7.5 Baseline multilevel models 2005 cohort

	English – null model – coefficient (standard error)	English – full model – coefficient (standard error)	Maths – null model – coefficient (standard error)	Maths – full model – coefficient (standard error)
Intercept	9.46 (3.86)	12.3 (5.38)	6.1 (2.7)	5.89 (3.8)
Federated?		NS		NS
Gender		NS		NS
Age		NS		NS
FSM		NS		NS
SEN		NS		NS
School size		NS		NS
Level 2 percentage variance	11.7	11.2	12.3	12.1
Level 1 percentage variance	88.3	88.8	87.7	87.9
Explained percentage variance level 2		5.3%		3.2%
Explained percentage variance level 1		0.0%		0.0%
Total percentage explained variance		0.6%		0.2%

NS = Variable not significant.

Table 7.6 2006 Multilevel models 2005 cohort

	English – null model – coefficient (standard error)	English – full model – coefficient (standard error)	Maths – null model – coefficient (standard error)	Maths – full model – coefficient (standard error)
Intercept	5.1 (1.3)	12.3 (5.38)	5.1 (0.4)	10.2 (3.8)
Federated?		NS		2.5 (1.1)
Gender		NS		NS
Age		NS		NS
FSM		NS		NS
SEN		NS		NS
School size		NS		NS
Level 2 percentage variance	10.8	10.2	12.8	9.5
Level 1 percentage variance	89.2	89.8	87.2	92.5
Explained percentage variance level 2		4.7%		36.5%
Explained percentage variance level 1		0.0%		1.2%
Total percentage explained variance		0.4%		4.7%

NS = Variable not significant.

Table 7.7 2007 Multilevel models 2005 cohort

	English – null model – coefficient (standard error)	English – full model – coefficient (standard error)	Maths – null model – coefficient (standard error)	Maths – full model – coefficient (standard error)
Intercept	6.4 (1.1)	5.6 (1.0)	5.1 (0.4)	10.2 (3.8)
Federated?		3.1 (1.4)		3.8 (1.1)
Gender		NS		NS
Age		NS		NS
FSM		NS		NS
SEN		NS		NS
School size		NS		NS
Level 2 percentage variance	14.9	9.4%	12.8	7.5
Level 1 percentage variance	85.1	90.6	87.2	92.5
Explained percentage variance level 2		38.9%		46.5%
Explained percentage variance level 1		0.0%		1.2%
Total percentage explained variance		6.1%		6.6%

NS = Variable not significant.

Table 7.8 Baseline multilevel measures 2006 cohort primary

	English – null model – coefficient (standard error)	English – full model – coefficient (standard error)	Maths – null model – coefficient (standard error)	Maths – full model – coefficient (standard error)
Intercept	4.2 (0.03)	4.6 (0.09)	4.1 (0.4)	4.3 (0.1)
Federated?		NS		NS
Gender		NS		NS
Age		NS		NS
FSM		NS		NS
SEN		–0.015 (0.005)		–0.009 (0.004)
School size		NS		NS
Level 2 percentage variance	12.8	10.2	15.4	13.9
Level 1 percentage variance	87.2	89.8	84.6	86.1
Explained percentage variance level 2		4.7%		5.2%
Explained percentage variance level 1		0.0%		0.2%
Total percentage explained variance		0.4%		0.7%

NS = Variable not significant.

and Maths is explained at the pupil level (level 1). However, variance at the school level is also significant. It is important here to point out that pupil level variance is not the same thing as pupil social background, as is often wrongly supposed. Rather, this may be a range of factors, including ability, motivation, and, to a large extent, measurement error.

As the samples were carefully matched on these variables, it is not surprising that most predictors were not significantly related to the outcomes. Federation is significantly related to outcomes in Maths in 2006 and 2007, and to outcomes in English in 2007. This is suggestive of impact, although other factors, such as prior capacity to change in federation as opposed to non-federation schools may of course be a causal factor as well. The impact of federation is quite strong in 2007, explaining nearly half of school level variance in Maths, and over a third of school level variance in English, making it a highly significant factor.

7.5 2006 Cohort

(i) Primary (Tables 7.8, 7.9, 7.10, and 7.11)

For cohort 2006 we can again see some evidence of impact of federations over time. Overall, the majority of the variance in both English and Maths is explained at the pupil level (level 1). However, variance at the school level is also significant and slightly larger for this cohort than for the 2005 cohort.

Table 7.9 2007 Multilevel models 2006 cohort primary

	English – null model – coefficient (standard error)	English – full model – coefficient (standard error)	Maths – null model – coefficient (standard error)	Maths – full model – coefficient (standard error)
Intercept	7.6 (1.3)	8.7 (0.7)	5.3 (0.6)	6.5 (0.8)
Federated?		2.4 (0.9)		3.1 (1.1)
Gender		NS		NS
Age		NS		NS
FSM		NS		NS
SEN		0.02 (0.01)		0.016 (0.007)
Number of heads		NS		NS
Level 2 percentage variance	13.2%	11.1%	15.8	12.7
Level 1 percentage variance	86.8	88.9	84.2	87.3
Explained percentage variance level 2		19.9%		22.7%
Explained percentage variance level 1		0.0%		0.%
Total percentage explained variance		2.3%		2.8%

NS = Variable not significant.

The variables on which the samples were matched were in general not significantly related to the outcomes. However, there was a weak significant relationship between percentage pupils with SEN and outcomes. Federation is significantly related to outcomes in English and Maths in 2007, and not at baseline. This is suggestive of impact, as again there appears to be an increase in impact over time. The impact of federation is quite strong in 2007, explaining nearly around 20% of the variance in outcomes.

Table 7.10 Baseline multilevel models 2006 cohort secondary

	Average points score at GCSE – null model	Average points score at GCSE – full model
Intercept	10.42 (0.22)	12.0 (0.94)
Federated?		NS
Gender		NS
Age		NS
FSM		NS
SEN		NS
School size		NS
Level 2 percentage variance	17.6%	16.2
Level 1 percentage variance	82.4	83.8
Explained percentage variance level 2		3.8%
Explained percentage variance level 1		0.2%
Total percentage explained variance		0.7%

NS = Variable not significant.

Table 7.11 2007 Multilevel models 2006 cohort secondary

	Average points score at GCSE – null model	Average points score at GCSE – full model
Intercept	7.6 (1.3)	8.0 (0.6)
Federated?		5.4 (2.6)
Gender		NS
Age		NS
FSM		NS
SEN		NS
Number of heads		NS
Level 2 percentage variance	19.0%	15.7%
Level 1 percentage variance	81.0	84.3
Explained percentage variance level 2		20.5%
Explained percentage variance level 1		0.0%
Total percentage explained variance		3.9%

NS = Variable not significant.

(ii) Secondary

Similar results are found for GCSE. Overall, the majority of the variance in both English and Maths is explained at the pupil level (level 1), with between 15 and 20% of the variance being at the school level.

As in the primary schools, most predictors were not significantly related to the outcomes. Federation is significantly related to outcomes in 2007, and not in 2006. This is suggestive of impact, as again there appears to be an increase in impact over time. The impact of federation is quite strong in 2007, explaining around 20% of the variance in outcomes.

7.5.1 Impact on Performance by Federation Type

In this section we will look at performance in the different types of federations. As sample sizes at level 2 (school level) are small in many cases, multilevel estimates may be unstable. Therefore we have used simple bivariate analyses to explore this question, which we would be able to interrogate in more detail if we had a larger sample of schools.

7.6 2005 Cohort

a. Cross-Phase Federation (Tables 7.12 and 7.13)

While no differences were found between federations and comparator schools at baseline, in 2006 and 2007 federation schools showed higher levels of performance than comparator schools.

Table 7.12 Cross-Phase federations – English

	Federation mean	Comparator mean	*t*
Baseline	4.1	3.8	−3.6
2006	4.2	3.6	−6.5**
2007	4.2	3.7	−5.7**

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Table 7.13 Cross-Phase federations – Maths

	Federation mean	Comparator mean	*t*
Baseline	4.0	3.7	−4.0
2006	4.2	3.6	−5.7**
2007	4.3	3.6	−5.9**

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Table 7.14 Size federations – English

	Federation mean	Comparator mean	*t*
Baseline	4.3	4.2	–1.0
2006	4.3	4.1	–1.2
2007	4.4	4.0	–2.1

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Table 7.15 Size federations – Maths

	Federation mean	Comparator mean	*t*
Baseline	4.2	4.2	0.0
2006	4.3	4.2	−0.7
2007	4.3	4.1	−1.1

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

b. Size Federations (Tables 7.14 and 7.15)

No significant differences were found for size federations.

Performance, Mainstreaming, and Faith federations were too few in number in this cohort for us to conduct analyses.

The tables above show that there is evidence that Cross-Phase federations may have a positive impact on performance, in that a federation schools in this category outperform comparison schools in years 2006 and 2007, but not 2005, but there is no evidence for size federations. The sample size for size federations was smaller, however.

7.7 2006 Cohort

(i) Primary

a. Cross-Phase Federation (Tables 7.16 and 7.17)

Table 7.16 Cross-Phase federations – English

	Federation mean	Comparator mean	*t*
Baseline	4.1	4.1	0.0
2007	4.1	4.0	2.7

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Table 7.17 Cross-Phase federations – Maths

	Federation mean	Comparator mean	*t*
Baseline	4.0	4.0	0.1
2007	4.0	4.0	0.3

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

No significant differences were found for Cross-Phase federations in primary for the 2006 cohort.

b. Performance Federations (Tables 7.18 and 7.19)

Table 7.18 Performance federations – English

	Federation mean	Comparator mean	*t*
Baseline	4.1	4.1	0.1
2007	4.2	3.8	2.9**

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

While no differences were found at baseline, in 2007 attainment in performance federations was significantly higher than in comparator schools in 2007.

Table 7.19 Performance federations – Maths

	Federation mean	Comparator mean	*T*
Baseline	4.1	4.2	1.0
2007	4.3	3.7	4.1***

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

c. Size Federations (Tables 7.20 and 7.21)

Table 7.20 Size federations – English

	Federation mean	Comparator mean	*t*
Baseline	4.1	4.4	–2.5*
2007	4.0	4.2	–1.5

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Table 7.21 Size federations – Maths

	Federation mean	Comparator mean	*t*
Baseline	4.1	4.3	–1.2
2007	4.0	4.1	–0.8

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Comparator schools showed higher performance than federation schools in English at baseline. No other significant differences were found.

d. Faith Federations (Tables 7.22 and 7.23)

Table 7.22 Faith federations – English

	Federation mean	Comparator mean	*t*
Baseline	4.1	4.1	−0.2
2007	4.1	3.9	1.2

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Table 7.23 Faith federations – Maths

	Federation mean	Comparator mean	*t*
Baseline	4.1	4.0	0.7
2007	4.0	3.9	0.5

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

No significant differences were found for Faith federations.

Mainstreaming federations were too few in number in this cohort for us to conduct analyses.

The only significant differences found were for Performance federations in 2007 in both English and Maths, where students outperformed their counterparts in the comparison schools (this had not been the case at baseline), and for Size federations in English at baseline, where comparator schools did better than federation schools. This was no longer the case in 2007.

Overall, it would appear that the main differences in performance between federation and comparator schools appear in Performance federations. The evidence for Cross-Phase federations is mixed, while few or no significant differences were found for the other types. It has to be pointed out though that in many cases sample sizes were too small to include particular federation types in the analyses.

(ii) Secondary

a. Cross-Phase Federation (Table 7.24)

Table 7.24 Cross-Phase federations

	Federation mean	Comparator mean	*t*
Baseline	341.5	351.8	–3.8***
2007	341.2	353.9	–2.6**

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

b. Performance Federation (Table 7.25)

Table 7.25 Performance federations

	Federation mean	Comparator mean	*t*
Baseline	295.6	274.8	1.9
2007	324.9	251.4	12.3***

* = Significant at the 0.05 level, ** = significant at the 0.01 level, *** = significant at the 0.001 level.

Notably in the secondary sample only Cross-Phase and Performance federations were present in sufficient numbers for analysis. In Cross-Phase federations, Comparator schools showed significantly higher levels of performance in both years. For Performance federations, there was a non-significant advantage for Performance schools at baseline, and a highly significant advantage for Performance schools in 2007.

Overall, it would appear that the main differences in performance between federation and comparator schools appear in Performance federations. The evidence for Cross-Phase federations is mixed, while few or no significant differences were found for the other types. It has to be pointed out though that in many cases sample sizes were too small to include particular federation types in the analyses.

7.7.1 Relationship with Ofsted Grades

In this section we will explore the extent to which there is a relationship between federation and Ofsted inspection grades. In view of sample size issues all types and phases of schools have been combined. It has to be pointed out here that as inspection does not occur on an annual basis in each school comparisons refer to different schools in different years, so any findings have to be considered indicative only (Tables 7.26, 7.27, and 7.28).

Table 7.26 2005 Inspections

Variable	Federated	Comparison	*t*
Overall effectiveness of provision	4.8	4.6	−0.6
Quality of teaching	3.6	3.5	−0.2
How well do learners achieve?	3.3	3.2	−0.1
Overall effectiveness of leadership and management	3.0	3.3	1.0

Table 7.27 2006 Inspections

Variable	Federated	Comparison	*t*
Overall effectiveness of provision	3.5	3.6	−0.1
Quality of teaching	2.4	2.4	0.1
How well do learners achieve?	2.5	2.5	0.2
Overall effectiveness of leadership and management	2.3	2.4	0.7

Table 7.28 2007 Inspections

Variable	Federated	Comparison	*t*
Overall effectiveness of provision	2.5	2.5	0.2
Quality of teaching	2.6	2.5	0.4
How well do learners achieve?	2.3	2.4	0.6
Overall effectiveness of leadership and management	2.1	2.1	−0.1

No significant differences were found between federation and comparison schools in any of the comparisons made for inspection ratings. However, it has to be pointed out that sample sizes were small.

7.7.2 Summary of Results

In conclusion we summarise the five key findings.

1. This study has identified six broad and sometimes overlapping categories of federation:

   - Size Federations
   - Cross-Phase Federations
   - Performance Federations
   - Faith Federations
   - Mainstreaming Federations
   - Academy groups

   The most popular category of federation in the sample is Cross-Phase federation and the least popular category is Academy groups.
2. There is evidence of impact on overall performance, in that while federation and comparator schools perform similarly at baseline, Federation is positively related to performance in the years following federation.
3. There is evidence to suggest that impact is strongest in Performance federations.
4. There is no relationship between federation and Ofsted judgements (grades).
5. There is no evidence of differential impact on students from different socio-economic settings, differences in gender, or with special educational needs.

7.8 Conclusion

Our analysis leads us to conclude that federations can have a positive impact on student outcomes and the federation impact is strongest where the aim of the federation is to raise educational standards by federating higher and lower attaining schools. This initial analysis would suggest persisting with the policy of federating schools to raise standards is a worthwhile enterprise.

However, if federations are to continue to be used as a structural solution we would draw attention to three major challenges within the system:

1. The challenge of stimulating and developing collaboration both within and between schools in very challenging contexts.
2. The challenge of developing appropriate accountability systems that move beyond single institutions as the primary unit of analysis.
3. The challenge of inspiring localised context specific approaches to improvement within an overarching national framework of intervention, such as the National Challenge.

We suggest these challenges need further exploration, including discussions with performance federation leaders to draw out the key issues related to the three challenges and the facilitators and barriers experienced while establishing a federation.