Data-driven analysis of factors affecting headline secondary school performance measures in England

Ian Widdows, Denis Newman-Griffis and Jo Bates
University of Sheffield, Information School

Context and motivation

- Secondary school performance measures (SPMs) introduced in England 32 years ago
- Today they affect schools' policy, funding, and choice
- Subject of numerous critiques and criticisms
- School accountability system is currently under scrutiny (e.g. UK general election, the aftermath of the death of headteacher Ruth Perry and the impact of COVID-19 on SPMs)
- We set out to use empirical data and a range of digital quantitative approaches to explore the nature, extent and breadth of the concept of '*performance*' represented by SPMs

Methods

- **Exploratory Factor Analysis (EFA)**
  to identify any latent constructs (or 'factors') underlying the set of metrics
- **Confirmatory Factor Analysis (CFA)**
  to test and confirm the underlying factor structure identified using EFA
- **Structural Equation Modelling (SEM)**
  to develop and examine more complex models
- **Correlation Analysis**
  to examine the strength of the associations between the 6 headline SPMs
- **Cluster Analysis**
  to explore the influence of school information metrics upon SPMs

Data

- Data from UK Government website "Compare the performance of schools and colleges in England"
- 3125 mainstream state secondary schools in England, 2021-2022.

Attainment 8
Progress 8
% grade 5+ in English and maths ('Basics')
%EBacc entry
EBacc Average Point Score
Pupil destinations

Table 1. The 6 headline SPMs

KS2 Average Point Score
% of pupils with Free School Meals ever in last six years
% English as additional language
% Special educational needs
school admission policy (selective or non-selective)
Whether the school has a religious character
School gender of entry (mixed or single gender)

Table 2. The 7 wider school information metrics

What we found

- The 6 headline SPMs have 4 underlying latent constructs which shape the concept of '*performance*'. These are: **'attainment', 'progress', 'curriculum' and 'destinations'**
- 3 out of the 6 headline SPMs measure the **same** factor i.e. '*attainment*'
- SPMs **do not** simply represent school '*performance*' factors – but are also shaped by the '*character*' and '*context*' of the school
- '*Attainment*' is particularly **strongly** influenced by school '*character*' and '*context*'
- Digital technologies (eg cluster analysis) may provide further insight into the challenges (or advantages) facing schools regarding SPMs
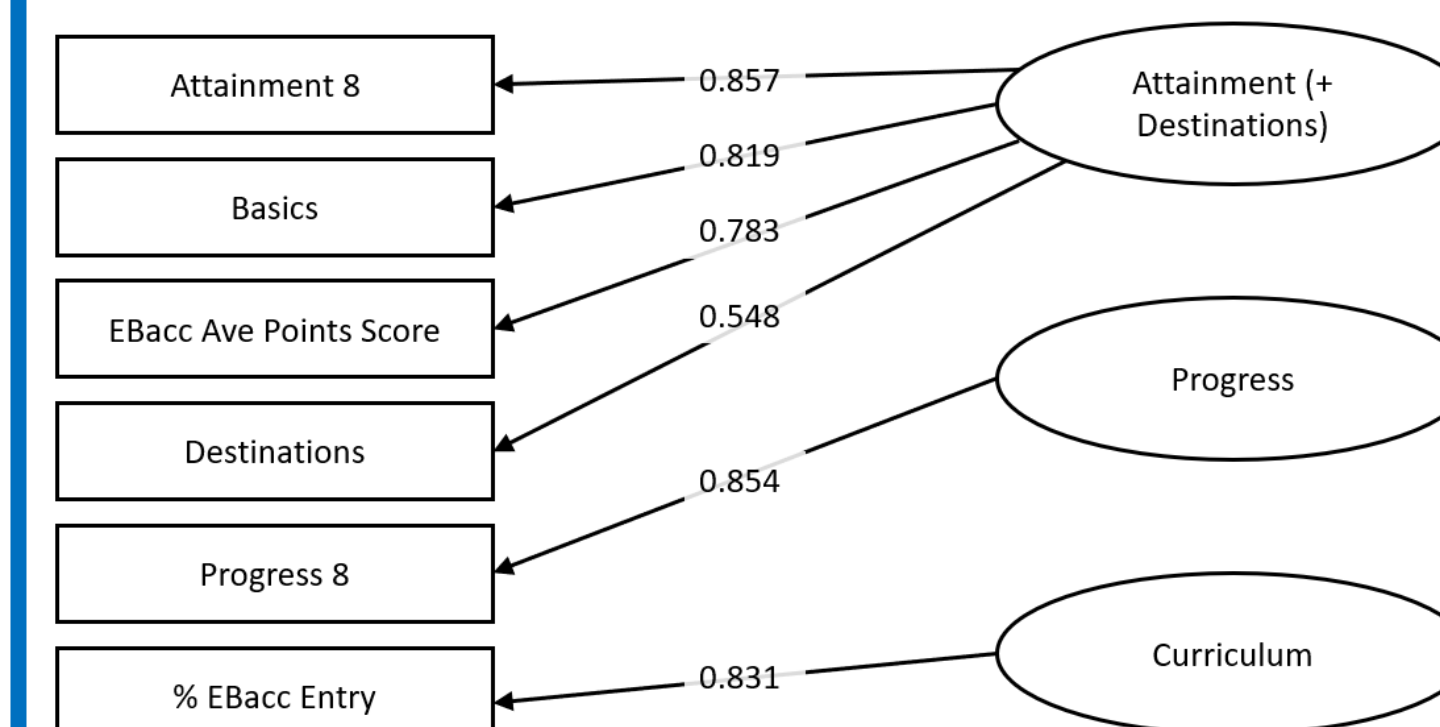

Figure 1. Factor Analysis Diagram showing factor loadings for 3 factor EFA for the 6 headline SPM variables

Using **EFA** we identified the latent factors which underpin SPMs: **'attainment', 'progress', 'curriculum' and 'destinations'**

Figure 2. Final model incorporating latent variables of school '*character*' and '*context*' and all four SPM latent variables

3 '*attainment*' SPMs are strongly influenced by school '*character*' and '*context*'

We used **CFA** to conceptualise the latent variables of school '*character*' and '*context*'.

Then using **SEM** we developed a series of more complex models, using a '*model, evaluate, refine*' approach
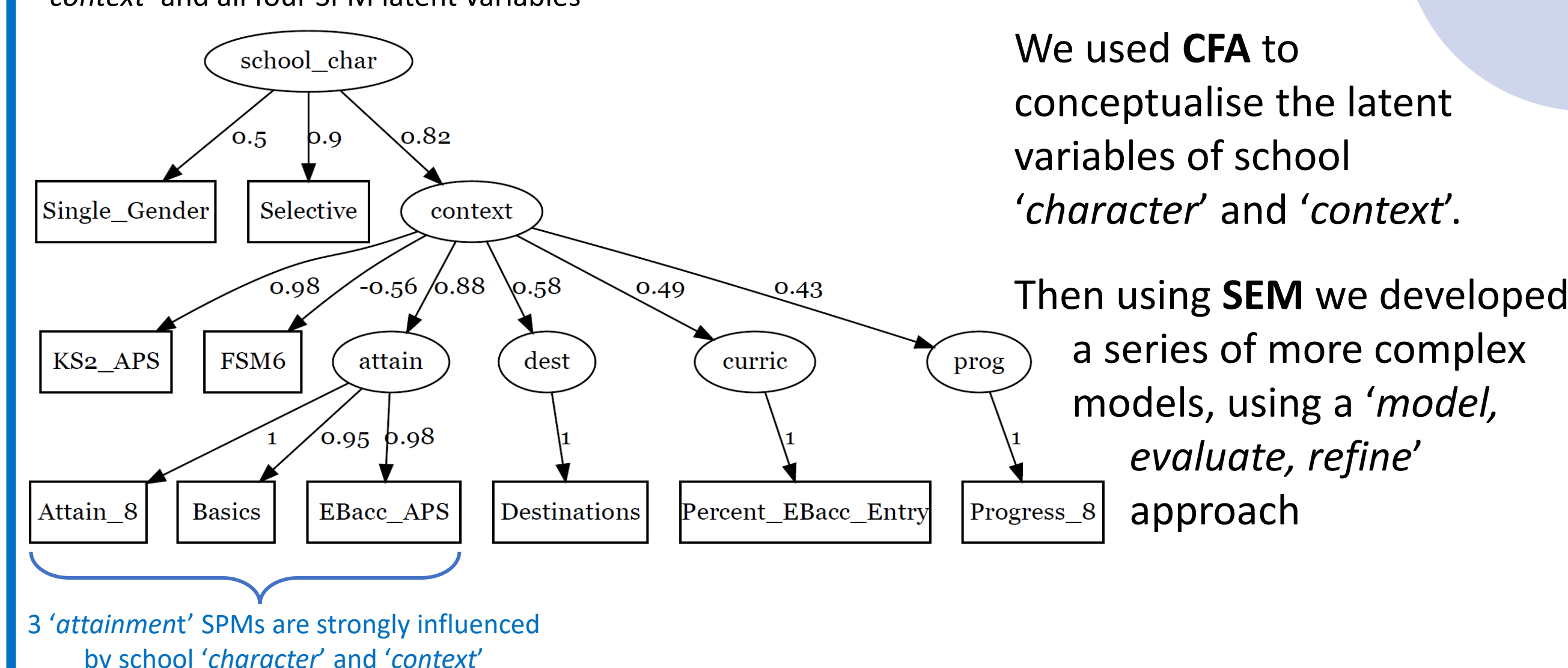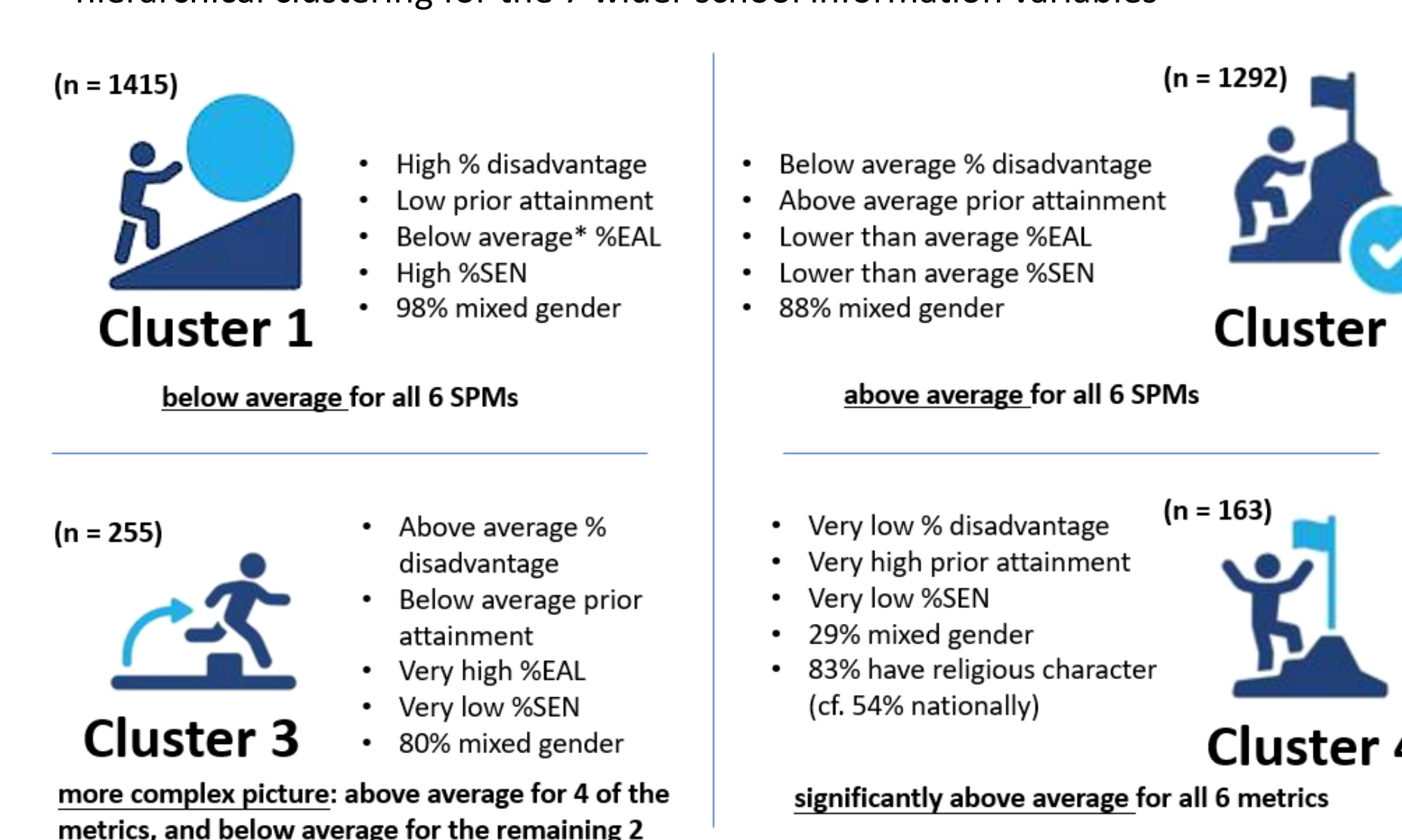
Discussion

Correlation Analysis showed us the strength of the correlation between the three '*attainment*' headline SPMs.

This suggests a significant narrowing of the concept of '*performance*' within the 6 headline SPMs.

Headline SPMs		Correlation
Attainment 8	EBacc Average Points Score	0.974
Attainment 8	% grade 5+ in English and maths ('Basics')	0.947
% grade 5+ in English and maths ('Basics')	EBacc Average Points Score	0.936

Table 3. Correlation between Attainment 8, EBacc APS and 'Basics' SPMs

Cluster Analysis showed potential in identifying and quantifying the challenges facing schools with regard to SPMs

Figure 3. Features of the 4 school clusters created using agglomerative hierarchical clustering for the 7 wider school information variables

(n = 1415)
- High % disadvantage
- Low prior attainment
- Below average* %EAL
- High %SEN
- 98% mixed gender

Cluster 1
below average for all 6 SPMs

(n = 1292)
- Below average % disadvantage
- Above average prior attainment
- Lower than average %EAL
- Lower than average %SEN
- 88% mixed gender

Cluster 2
above average for all 6 SPMs

(n = 255)
- Above average % disadvantage
- Below average prior attainment
- Very high %EAL
- Very low %SEN
- 80% mixed gender

Cluster 3
more complex picture: above average for 4 of the metrics, and below average for the remaining 2

(n = 163)
- Very low % disadvantage
- Very high prior attainment
- Very low %SEN
- 29% mixed gender
- 83% have religious character (cf. 54% nationally)

Cluster 4
significantly above average for all 6 metrics

[*national average for all schools]

Questions? Comments?

icwiddows1@sheffield.ac.uk

icwiddows.co.uk

@wids_widdows

linkedin.com/in/ian-c-widdows-33a767247

Implications and next steps

- The dominance of '*attainment*' in SPMs has significant implications for their use (informing parental choice, inspection, intervention and school improvement planning)
- There is an opportunity to replace 2 of SPMs with other metrics, adding breadth to SPMs
- Influence of school '*character*' and '*context*' (factors which are not school '*performance*' related) on SPMs has important implications for both their design and interpretation
- Use of digital analytical approaches may provide a mechanism of identifying and quantifying the challenges facing schools based on their '*character*' and '*context*'.
- Potential to inform policy around support for schools in overcoming such challenges

This work will be extended in a number of ways:

- Larger datasets will be used (National Pupil Database and School Workforce Census) - a greater range of 'wider school information' variables and allowing analysis at more granular student level, not just aggregated at school level
- SEM models developed here will be developed further using the same '*model, evaluate, refine*' approach using subsets of the data, e.g. schools with similar inspection gradings or within geographical regions, and with larger datasets
- Alternative clustering approaches will be used and evaluated to further explore the potential of such analyses in the design of SPMs and related policy