Communicating uncertainty in school league tables

George Leckie and Harvey Goldstein
Centre for Multilevel Modelling
University of Bristol
A brief history of England’s school league tables

• 1994 onwards: %5+A*-C GCSE grades
  – Unfair since schools differ in the quality of their intakes
  – Frequently misinterpreted as a measure of school quality

• 2006 onwards: **Contextual value-added (CVA) scores**
  – Adjusts for the intake achievement of students
  – Multilevel model based estimates of school quality
  – Published with 95% confidence intervals

• Justification: **School accountability** and **school choice**
The CVA model is a two-level multilevel model

- A simplified version of this model can be written as:

\[ y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij} \]

\[ u_j \sim N(0, \sigma_u^2) \]

\[ e_{ij} \sim N(0, \sigma_e^2) \]

- \( y_{ij} \) is the age 16 GCSE score for student \( i \) in school \( j \)
- \( x_{ij} \) is their age 11 Key Stage 2 score
- \( u_j \) is the school level random effect for secondary school \( j \)
- \( e_{ij} \) is the student level random effect or residual
The CVA scores and their 95% confidence intervals

- The CVA scores are the posterior estimates of the school effects

\[ \hat{u}_j = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_e^2}{n_j}} (y_{ij} - \hat{y}_{ij}) \]

- Their associated variances are given by

\[ \text{var}(\hat{u}_j - u_j) = \frac{\sigma_u^2 \sigma_e^2}{n_j \sigma_u^2 + \sigma_e^2} \]

- The 95% confidence intervals for the CVA scores are calculated as

\[ \left( \hat{u}_j - 1.96 \sqrt{\text{var}(\hat{u}_j - u_j)}, \hat{u}_j + 1.96 \sqrt{\text{var}(\hat{u}_j - u_j)} \right) \]
The media rank schools and ignore uncertainty

• The media ignore the magnitude of the CVA scores and simply rank schools
  – CVA scale/metric is difficult to understand

• The media ignore uncertainty; they present CVA scores as if they were perfect measurements of school quality
  – 95% confidence intervals are difficult to understand

• The result is that users are encouraged to over interpret even the smallest differences in schools’ ranks as genuine differences in school quality
  – Users are likely to make misleading inferences
  – Users may place too much importance on school league tables
What we do

• We present a simulation method to communicate the uncertainty in CVA scores in terms of the probability, chance, or odds, that each school is the best rather than in terms of 95% confidence intervals.

• We describe our approach in the context of a parent using the CVA school league tables to choose which school to send their child to.
## Local Authority: Bristol, City of

<table>
<thead>
<tr>
<th>Institution Name</th>
<th>Measure centred around 1000</th>
<th>Limit of Key Stage 2 to 4 CVA Confidence Intervals</th>
<th>Coverage</th>
<th>Number of qualifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Upper</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Local Authority Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>England (state funded schools only)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>England (all schools)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashton Park School - Bristol S</td>
<td>1003.2</td>
<td>1012.8</td>
<td>993.7</td>
<td>98%</td>
</tr>
<tr>
<td>Bedminster Down School - Bristol</td>
<td>988.5</td>
<td>997.8</td>
<td>979.3</td>
<td>98%</td>
</tr>
<tr>
<td>Bridge Learning Campus - Secondary - Bristol</td>
<td>1003.4</td>
<td>1013.9</td>
<td>992.8</td>
<td>98%</td>
</tr>
<tr>
<td>Brislington Enterprise College - Bristol S</td>
<td>970.3</td>
<td>979.1</td>
<td>961.5</td>
<td>97%</td>
</tr>
<tr>
<td>Bristol Brunel Academy - Bristol</td>
<td>1005.0</td>
<td>1015.8</td>
<td>994.2</td>
<td>85%</td>
</tr>
<tr>
<td>Bristol Cathedral Choir School - Bristol</td>
<td>1002.0</td>
<td>1017.0</td>
<td>987.0</td>
<td>87%</td>
</tr>
</tbody>
</table>
## Secondary schools and colleges in Bristol

<table>
<thead>
<tr>
<th>School name</th>
<th>% 5 A*-C GCSEs</th>
<th>English Bacc</th>
<th>CVA</th>
<th>A/AS points</th>
</tr>
</thead>
<tbody>
<tr>
<td>The City Academy Bristol (AC)</td>
<td>36</td>
<td>3</td>
<td>1036.2</td>
<td>592</td>
</tr>
<tr>
<td>Oasis Academy Brightstowe (AC)</td>
<td>29</td>
<td>1</td>
<td>1028.2</td>
<td>NA</td>
</tr>
<tr>
<td>Cotham School</td>
<td>77</td>
<td>41</td>
<td>1015.7</td>
<td>825.7</td>
</tr>
<tr>
<td>St Mary Redcliffe and Temple School</td>
<td>70</td>
<td>21</td>
<td>1013.4</td>
<td>796.1</td>
</tr>
<tr>
<td>Bristol Metropolitan Academy (AC)</td>
<td>39</td>
<td>7</td>
<td>1010.8</td>
<td>NA</td>
</tr>
<tr>
<td>Colston's Girls' School (AC)</td>
<td>91</td>
<td>54</td>
<td>1009.6</td>
<td>706.6</td>
</tr>
<tr>
<td>Merchants' Academy (AC)</td>
<td>25</td>
<td>1</td>
<td>1009.6</td>
<td>SS</td>
</tr>
<tr>
<td>Oasis Academy Bristol (AC)</td>
<td>29</td>
<td>0</td>
<td>1007.4</td>
<td>NA</td>
</tr>
<tr>
<td>St Bede's Catholic College</td>
<td>72</td>
<td>28</td>
<td>1006</td>
<td>NA</td>
</tr>
<tr>
<td>Bristol Brunel Academy (AC)</td>
<td>45</td>
<td>4</td>
<td>1005</td>
<td>422.3</td>
</tr>
<tr>
<td>Orchard School</td>
<td>37</td>
<td>5</td>
<td>1004.7</td>
<td>NA</td>
</tr>
<tr>
<td>Fairfield High School</td>
<td>49</td>
<td>19</td>
<td>1004.2</td>
<td>NA</td>
</tr>
</tbody>
</table>
CVA scores with 95% confidence intervals

- 12 out of 19 schools are not significantly different from the national average
- Many schools will also not be significantly different from one another
Where shall I send my child to school?

- I have to inform Bristol LA my top three preferences
- Let’s choose between the three nearest schools
Bristol LA common application form for transfer to secondary education in September 2011

Preferences

Please enter the names, in priority order of any maintained secondary school(s) (including Voluntary Aided) that you would like your child to attend. Whilst all admission authorities will have regard to the reasons for your school preferences, giving reasons for your preference does not guarantee a place at your preferred school or mean that admission authorities can deviate from their published admissions policies.

1st Preference School

My 1st preference is: ___________________________________________ School

Reasons for preference/additional information

Medical/social reasons: ☐ Other: ___________________________________________

Sibling already at the school (please give details):

Forename: ___________________________ Middle name/s: ___________________________
Surname: ___________________________
Date of birth: Day __ Mth __ Year _____ Gender: male ☐ female ☐
Present school: ___________________________________________

Home address (if different to address(es) given for the parent/carer or the child for whom application is being made):

________________________________________________________________________
________________________________________________________________________
• How certain are we that this is the correct ordering of schools?

• The less certain we are, the less weight we should place on CVA as a guide to school choice
Implied probability that each school is the 1\textsuperscript{st}, 2\textsuperscript{nd} or 3\textsuperscript{rd} best

- The media's presentation of CVA implies that we are completely certain that the blue school is the best, then the red school and then the green school.
CVA scores with 95% confidence intervals

- The three schools’ confidence intervals overlap suggesting, that the true ordering of schools is by no means certain
Implied normal sampling distributions for each school

- The normal sampling distributions can be viewed as summaries of the relative plausibility of each possible CVA score for the true CVA score for each school.
Simulation method

1. Draw one score at random from each “probability” distribution
2. Rank the three scores
3. Repeat 10,000 times and summarise the results
Iteration 1

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Blue school</th>
<th>Red school</th>
<th>Green school</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
</tr>
</tbody>
</table>
Iteration 2

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Blue school</th>
<th>Red school</th>
<th>Green school</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1st</td>
<td>3rd</td>
<td>2nd</td>
</tr>
</tbody>
</table>
Iteration 10,000

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Blue school</th>
<th>Red school</th>
<th>Green school</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,000</td>
<td>2nd</td>
<td>1st</td>
<td>3rd</td>
</tr>
</tbody>
</table>
## Simulated school league tables

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Blue school</th>
<th>Red school</th>
<th>Green school</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10,000</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

- The blue school came first on 6,907 of the 10,000 iterations
- The probability that the blue school is the best is 0.69
- The blue school has a 7 out of 10 chance of being the best
- The odds that the blue school is the best are 11 to 4
- So while the blue school looks the best, this is by no means guaranteed
Actual probability that each school is the 1\textsuperscript{st}, 2\textsuperscript{nd} or 3\textsuperscript{rd} best
Implied probability that each school is the 1\textsuperscript{st}, 2\textsuperscript{nd} or 3\textsuperscript{rd} best
Past performance is no guarantee of future performance ... in many cases the value of the investment can fall as well as rise.
Seven years out of date!

- The 2010 school league table reports the estimated effectiveness of each school for the **2010 GCSE cohort**

- However, parents choosing schools want to instead know how effective schools will be for the **2017 GCSE cohort**

- Schools’ CVA scores and league table positions change from year to year; Leckie and Goldstein (2009) report a **correlation of 0.64 for five-years-apart**

- Inferences about the future effectiveness of schools will be **far less precise** than inferences about their current effectiveness

- We must **factor in the additional uncertainty** that arises from predicting seven years into the future
CVA scores for the 2017 cohort

- Leckie and Goldstein (2009) derive formula for the CVA scores for a future cohort based on the CVA scores of the current cohort

\[
\hat{u}_j = 0.64 \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_e^2}{n_j}} (y_{ij} - \hat{y}_{ij})
\]

- Their associated variances are given by

\[
\text{var}(\hat{u}_j - u_j) = \frac{n_j \sigma_u^4(1 - 0.64^2) + \sigma_u^2 \sigma_e^2}{n_j \sigma_u^2 + \sigma_e^2}
\]

- The 95% confidence intervals for the CVA scores are calculated as

\[
\left(\hat{u}_j - 1.96 \sqrt{\text{var}(\hat{u}_j - u_j)}, \hat{u}_j + 1.96 \sqrt{\text{var}(\hat{u}_j - u_j)}\right)
\]
CVA scores with 95% confidence intervals (2017)

- This is the relevant information for parents choosing schools
CVA scores with 95% confidence intervals (2010)

- This is the relevant information for holding schools accountable
Actual probability that each school is the 1\textsuperscript{st}, 2\textsuperscript{nd} or 3\textsuperscript{rd} best (2017)

- The probability that the blue school will be the best for the 2017 cohort is 0.44
Actual probability that each school is the 1st, 2nd or 3rd best (2010)

- The probability that the blue school was the best for the 2010 cohort is 0.69
There are also many other limitations of the CVA model

• At GCSE, students take **different combinations of subjects**

• Schools will be **differentially effective** for different types of students and for different responses

• **Student mobility** between schools is not recognised

• Students with **missing data** are listwise deleted

• Little is known about the **inter-rater reliability** of the tests
Broader limitations of school league tables

• Huge financial cost to implement

• Teaching time is taken up with the administrative burden of the tests

• The range of knowledge and skills that tests assess is very narrow

• Stress caused by over-testing turns children off education
Boycott threatens Sats test chaos

Head teachers have announced plans to boycott next month's Sats tests for primary school pupils in England.

Confusion now surrounds the national tests due to be taken by hundreds of thousands of 10 and 11-year-olds.

Mick Brookes, head of the National Association of Head Teachers which is boycotting the tests, said it was targeting a "flawed testing regime".

Schools Secretary Ed Balls urged head teachers to "think hard before disrupting children's learning".

Uncertainty remains over how much disruption will be caused - with Mr Brookes saying there would be no "coercion" to persuade head teachers who did not want to take part in the boycott.

'Meaningless tables'

Mr Balls said there was a statutory duty on heads to oversee tests - and that he would issue advice to governors and local authorities on the threat of a boycott.

Head teachers say they will
Conclusions

• The issue of statistical uncertainty is a fundamental aspect of reporting CVA, but is ignored by the media

• One reason for this is that the 95% confidence intervals are hard to understand

• We presented a simulation based method to communicate uncertainty in terms of the probability, chance, or odds, that each school is the best

• However, it is important to realise that there are many other statistical and much broader criticisms of school league tables which also need to be addressed
Conclusions (cont.)

• Bespoke comparisons such as parents choosing between local schools seem to fit very naturally with our simulation method.

• However, for school choice purposes CVA suffers from a major limitation.

• The CVA school performance tables understate the substantial uncertainty in using current CVA as a guide to future CVA.

  – Ignoring this additional source of uncertainty misleads parents into believing that the league tables are more informative than they truly are.

  – Implies that parents should place less weight on CVA and more weight on other sources of information available to them.
References


