Predicting the Past
A Retrodictive Model for Modern Rugby Union

Ian Hamilton
University of Warwick

18th February 2020
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated?
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated?

A: Rugby union!
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin?
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin?

A: Schools rugby!
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches?
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches?

A: Daily Mail Trophy!
Motivation

Q: Wouldn’t it be nice (for me, at least) if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches, and the methodology they currently use could do with some serious improvement?
Motivation

Q: Wouldn’t it be nice if there was a sport with which I was familiar, where the points system was just a bit more complicated, where there was a system of matches that do not make up a full round robin, and there was an actual tournament based on the results of these matches, and the methodology they currently use could do with some serious improvement?

A: Full house!
Bradley Terry

In the context of tournaments, the probability that team $i$ beats team $j$ is given by

$$P(i \succ j) = \frac{\pi_i}{\pi_i + \pi_j}$$

where $\pi_i$ is positive-valued, and can be thought of as a parameter reflecting the strength of team $i$.

Zermelo (1929), Bradley & Terry (1952)
Extension to include ties

\[ P(i \succ j) = \frac{\pi_i}{\pi_i + \pi_j + \nu \sqrt{\pi_i \pi_j}} \]

\[ P(i \approx j) = \frac{\nu \sqrt{\pi_i \pi_j}}{\pi_i + \pi_j + \nu \sqrt{\pi_i \pi_j}} \]

Davidson (1970)
Extension to account for home advantage (order effects)

\[ P(i \succ j) = \frac{\pi_i}{\pi_i + \gamma \pi_j + \nu \sqrt{\pi_i \pi_j}} \]

\[ P(i \prec j) = \frac{\gamma \pi_j}{\pi_i + \gamma \pi_j + \nu \sqrt{\pi_i \pi_j}} \]

\[ P(i \approx j) = \frac{\nu \sqrt{\pi_i \pi_j}}{\pi_i + \gamma \pi_j + \nu \sqrt{\pi_i \pi_j}} \]

Davidson & Beaver (1977)
Applying to 3 for a win, 1 for a draw

\[ P(i \succ j) = \frac{\pi_i}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}} \]

\[ P(i \approx j) = \frac{\nu(\pi_i \pi_j)^{\frac{1}{3}}}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}} \]

See: alt-3.uk

Firth (2017)
Rugby union scoring rule

League Points:

4 points for a win
2 points for a draw
0 points for a loss
1 bonus point for losing by less than seven points
1 bonus point for scoring four or more tries
<table>
<thead>
<tr>
<th>Model</th>
<th>B-T</th>
<th>Davidson</th>
<th>Firth</th>
<th>Rugby</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points - win</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Points - draw</td>
<td>NA</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Points - other</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1 (try, losing)</td>
</tr>
<tr>
<td>Model - i win</td>
<td>$\pi_i$</td>
<td>$\pi_i$</td>
<td>$\pi_i$</td>
<td>???</td>
</tr>
<tr>
<td>Model - draw</td>
<td>NA</td>
<td>$(\pi_i \pi_j)^{1/2}$</td>
<td>$(\pi_i \pi_j)^{1/3}$</td>
<td>???</td>
</tr>
<tr>
<td>Model - other</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>???</td>
</tr>
</tbody>
</table>
RASR (pronounced ‘razor’) - Ranking Algorithm for Schools Rugby

Part one: result outcome

\[ P(\text{team } i \text{ beats team } j \text{ by wide margin}) \propto \tau_i^4 \pi_i^4 \]
\[ P(\text{team } i \text{ beats team } j \text{ by narrow margin}) \propto \kappa \tau_i^3 \pi_i^4 \pi_j \]
\[ P(\text{team } i \text{ draws with team } j) \propto \nu \pi_i^2 \pi_j^2 \]
\[ P(\text{team } j \text{ beats team } i \text{ by narrow margin}) \propto \frac{\kappa \pi_i \pi_j^4}{\tau^3} \]
\[ P(\text{team } j \text{ beats team } i \text{ by wide margin}) \propto \frac{\pi_j^4}{\tau^4} \]
Part two: try bonus outcome

\[ P(\text{team } i \text{ and team } j \text{ both gain try bonus point}) \propto \theta \pi_i \pi_j \]

\[ P(\text{only team } i \text{ gains try bonus point}) \propto \tau \pi_i \]

\[ P(\text{only team } j \text{ gains try bonus point}) \propto \frac{\pi_j}{\tau} \]

\[ P(\text{neither team gains try bonus point}) \propto \phi \]
A principle-based approach

Maximise entropy

\[ S(p) = - \sum_{i,j} \sum_{a,b} p_{i,j}^{a,b} \log p_{i,j}^{a,b} , \]

subject to conditions,

\[ \sum_{a,b} p_{i,j}^{a,b} = 1 , \]  

(1)

and

\[ \sum_{j} \sum_{a,b} a p_{i,j}^{a,b} = \sum_{j} \sum_{a,b} a m_{i,j}^{a,b} , \]  

(2)

where \( p_{i,j}^{a,b} \) is the probability that \( i \) gains \( a \) points and \( j \) gains \( b \) points, and \( m_{i,j}^{a,b} \) is the number of matches that have resulted with \( i \) gaining \( a \) points and \( j \) gaining \( b \) points.
A principle-based approach

Taking the Lagrangian and differentiating wrt $p_{a,b}^{ij}$ we have

$$\log p_{a,b}^{ij} = -\lambda_{ij} - a\lambda_i - b\lambda_j - 1,$$

which gives us that

$$p_{a,b}^{ij} \propto \pi_i^a \pi_j^b,$$

where the $\pi_i = \exp(-\lambda_i)$, may be used to rank the teams, and $\exp(-\lambda_{ij} - 1)$ is the constant of proportionality.
Potential models

Examples:

- Try bonus dependent on result outcome and opposition
- Try bonus independent of result outcome but dependent on opposition
- Try bonus independent of result outcome and opposition
- Offensive-defensive strengths
- Home-away strengths
To prior or not to prior?

Introduce a dummy $team_0$ against whom each other team wins one and loses one, then decide how much weight to give these matches.

Pros:
- Ensures connectedness therefore rating from start of season
- Explicitly controls fairness in situations of varying fixture numbers
- Allows for estimation of structural parameters even with existence of 100% record

Cons:
- Might not match intuition / round robin outcomes
Intuitive Measure

Projected Points per Match

\[ \text{PPPM}_i = \frac{1}{n-1} \sum_j \sum_{a,b} a p_{ij} \]

Intuitive measure that converges to the rating in round robin
League Points per Match + Additional Points

Additional Points in the Daily Mail Trophy are awarded based on the ranking of the current season’s opponents in the previous season’s tournament:

- Rank 1 to 25: 0.3
- Rank 26 to 50: 0.2
- Rank 51 to 75: 0.1
- Otherwise: 0
<table>
<thead>
<tr>
<th>School</th>
<th>DMT Rank</th>
<th>DMT</th>
<th>PPPM Rank</th>
<th>PPPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wellington College</td>
<td>1</td>
<td>6.46</td>
<td>7</td>
<td>3.73</td>
</tr>
<tr>
<td>Kirkham</td>
<td>2</td>
<td>6.44</td>
<td>1</td>
<td>4.41</td>
</tr>
<tr>
<td>Bedford</td>
<td>3</td>
<td>6.35</td>
<td>2</td>
<td>4.37</td>
</tr>
<tr>
<td>Bromsgrove</td>
<td>4</td>
<td>6.21</td>
<td>4</td>
<td>4.15</td>
</tr>
<tr>
<td>Sedbergh</td>
<td>5</td>
<td>6.10</td>
<td>5</td>
<td>3.99</td>
</tr>
<tr>
<td>Woodhouse Grove</td>
<td>6</td>
<td>5.65</td>
<td>19</td>
<td>3.31</td>
</tr>
<tr>
<td>Millfield</td>
<td>7</td>
<td>5.21</td>
<td>13</td>
<td>3.64</td>
</tr>
<tr>
<td>Clifton College</td>
<td>8</td>
<td>5.11</td>
<td>8</td>
<td>3.73</td>
</tr>
<tr>
<td>Solihull</td>
<td>9</td>
<td>5.10</td>
<td>11</td>
<td>3.67</td>
</tr>
<tr>
<td>St Paul’s</td>
<td>9</td>
<td>5.10</td>
<td>14</td>
<td>3.58</td>
</tr>
</tbody>
</table>
## Results 2016/17

<table>
<thead>
<tr>
<th>School</th>
<th>DMT Rank</th>
<th>DMT</th>
<th>PPPM Rank</th>
<th>PPPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wellington College</td>
<td>1</td>
<td>7.22</td>
<td>3</td>
<td>4.37</td>
</tr>
<tr>
<td>Sedbergh</td>
<td>2</td>
<td>6.50</td>
<td>2</td>
<td>4.43</td>
</tr>
<tr>
<td>Harrow</td>
<td>3</td>
<td>6.34</td>
<td>6</td>
<td>4.22</td>
</tr>
<tr>
<td>St Peter’s, York</td>
<td>4</td>
<td>6.23</td>
<td>8</td>
<td>4.06</td>
</tr>
<tr>
<td>Kirkham</td>
<td>5</td>
<td>6.15</td>
<td>1</td>
<td>4.61</td>
</tr>
<tr>
<td>Canford</td>
<td>6</td>
<td>6.10</td>
<td>9</td>
<td>4.02</td>
</tr>
<tr>
<td>Clifton College</td>
<td>7</td>
<td>6.00</td>
<td>5</td>
<td>4.25</td>
</tr>
<tr>
<td>Rugby</td>
<td>8</td>
<td>5.96</td>
<td>7</td>
<td>4.06</td>
</tr>
<tr>
<td>Brighton College</td>
<td>9</td>
<td>5.90</td>
<td>4</td>
<td>4.29</td>
</tr>
<tr>
<td>Woodhouse Grove</td>
<td>10</td>
<td>5.81</td>
<td>12</td>
<td>3.93</td>
</tr>
</tbody>
</table>
# Results 2017/18

<table>
<thead>
<tr>
<th>School</th>
<th>DMT Rank</th>
<th>DMT</th>
<th>PPPM Rank</th>
<th>PPPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedbergh</td>
<td>1</td>
<td>7.41</td>
<td>1</td>
<td>4.65</td>
</tr>
<tr>
<td>Wellington College</td>
<td>2</td>
<td>7.18</td>
<td>7</td>
<td>4.18</td>
</tr>
<tr>
<td>Cranleigh</td>
<td>3</td>
<td>6.33</td>
<td>4</td>
<td>4.32</td>
</tr>
<tr>
<td>Harrow</td>
<td>4</td>
<td>6.20</td>
<td>3</td>
<td>4.33</td>
</tr>
<tr>
<td>Cheltenham College</td>
<td>5</td>
<td>6.16</td>
<td>8</td>
<td>4.07</td>
</tr>
<tr>
<td>St Peter’s, York</td>
<td>6</td>
<td>5.83</td>
<td>6</td>
<td>4.19</td>
</tr>
<tr>
<td>Brighton College</td>
<td>7</td>
<td>5.63</td>
<td>20</td>
<td>3.59</td>
</tr>
<tr>
<td>Reed’s</td>
<td>8</td>
<td>5.50</td>
<td>2</td>
<td>4.38</td>
</tr>
<tr>
<td>Clifton College</td>
<td>8</td>
<td>5.50</td>
<td>16</td>
<td>3.72</td>
</tr>
<tr>
<td>Haileybury</td>
<td>10</td>
<td>5.49</td>
<td>10</td>
<td>4.02</td>
</tr>
</tbody>
</table>