# RSS Euro 2020 Prediction Competition

## Bradley-Terry Go Home

Ian Hamilton, Stefan Stein, David Selby

8th September 2021

# Motivating ideas

1. A significant proportion of the competition outcome would be luck.

# Motivating ideas

1. A significant proportion of the competition outcome would be luck.
2. We don't have much time!

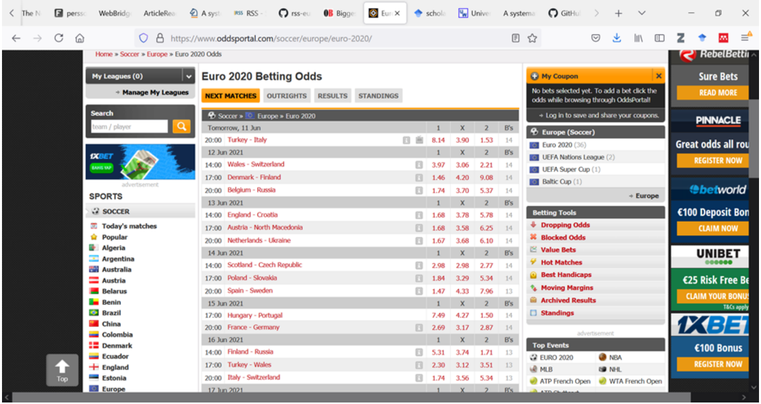
# Motivating ideas

1. A significant proportion of the competition outcome would be luck.
2. We don't have much time!
3. Markets are good at evaluation.

# Motivating ideas

1. A significant proportion of the competition outcome would be luck.
2. We don't have much time!
3. Markets are good at evaluation.
4. Markets get things wrong in (somewhat) predictable ways.

# The data - match probabilities



# The model - Bradley-Terry

Probability that $i$ beats $j$

$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j},$$

where $\pi_i$ is the 'strength' of $i$.

As a generalised linear model

$$\text{logit}(p_{ij}) = \lambda_i - \lambda_j,$$

where $\lambda_i = \log(\pi_i)$

Zermelo (1929); Bradley and Terry (1952)

# Bradley-Terry - typical use

Bradley-Terry model applied to a set of results, for the purpose of prediction or ranking e.g. alt-3.uk

Parameters estimated by maximum likelihood estimation

$$L(\boldsymbol{\lambda}) = \prod_{i<j} \binom{m_{ij}}{c_{ij}} p_{ij}^{c_{ij}} (1 - p_{ij})^{m_{ij} - c_{ij}},$$

where $c_{ij}$ is the number of time $i$ beats $j$ and $m_{ij} = c_{ij} + c_{ji}$ is the number of matches between $i$ and $j$.

# Bradley-Terry - issues

But:

1. not enough recent useful results to estimate strengths reliably
2. market prices are likely to be more informative
3. draws in the group stages

# Bradley-Terry - dealing with draws

Extension to draws (alt-3.uk, Davidson (1970))

$$\mathbb{P}(i \text{ beats } j) = \frac{\pi_i}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}}$$

$$\mathbb{P}(i \text{ draws with } j) = \frac{\nu(\pi_i \pi_j)^{\frac{1}{3}}}{\pi_i + \pi_j + \nu(\pi_i \pi_j)^{\frac{1}{3}}}$$

Note even with draws:

$$\frac{p_{ij}}{p_{ji}} = \frac{\pi_i}{\pi_j} \quad \text{or} \quad \text{logit}(p_{ij}) = \lambda_i - \lambda_j$$

# Intra-group strength estimation

Can estimate the intra-group log-strengths $r_i = \log s_i$ by linear regression:

$$\log\left(\frac{p_{ij}}{p_{ji}}\right) = r_i - r_j,$$

since $p_{ij}$ are known from market odds.

But how do we compare strengths between groups?

# Overall strength estimation

Assumptions:

1. Team $i$'s overall strength $\pi_i$ is a scaling of its intra-group strength $s_i$ by a factor dependent on its group $\gamma_{G(i)}$

$$\pi_i = \gamma_{G(i)} s_i \quad \text{or equivalently} \quad \lambda_i = \log \gamma_{G(i)} + r_i$$

2. The strength of every team's unknown final opponent is the same

$$p_{io} = \mathbb{P}(i \text{ winning tournament} \mid i \text{ reaches final}) = \frac{\pi_i}{\pi_i + \pi_o},$$

where $\pi_o$ is the strength of the unknown final opponent.

# Overall strength estimation

We can calculate $p_{io}$ from market odds since

$$p_{io} = \frac{\mathbb{P}(i \text{ winning tournament})}{\mathbb{P}(i \text{ reaches final})}$$

Then we have that

$$\log\left(\frac{p_{io}}{p_{oi}}\right) = \lambda_i - \lambda_o = \log \gamma_{G(i)} + r_i - \lambda_o,$$

and we can estimate $\log \gamma_{G(i)}$ and $\lambda_o$ through linear regression.

# Knock-out prediction

Now we can calculate the strengths of each team

$$\pi_i = \gamma_{G(i)} s_i,$$

and apply these through the Bradley-Terry model to predict the KO match results

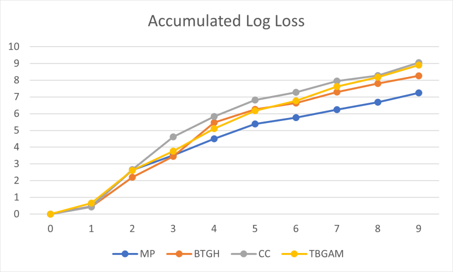
$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j}.$$

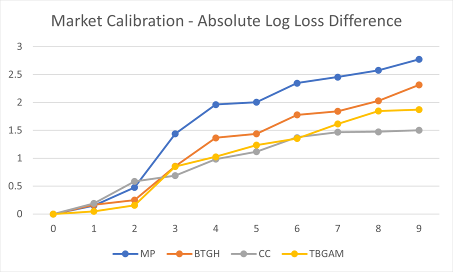
# Miscellaneous notes

1. Parsimonious model - 120 data points (72 group stage match probabilities + 24 reach final + 24 tournament win); two linear regressions; two days
2. What happened to market being wrong in predictable ways?
3. Did we do well just because of taking market odds for the group stage?

# How much was luck?

Performance graphs for KO stages alone based on the nine competition updates



# Bibliography


Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, 39(3/4):324–345.

Davidson, R. R. (1970). On extending the bradley-terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association*, 65(329):317–328.

Zermelo, E. (1929). Die berechnung der turnier-ergebnisse als ein maximumproblem der wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 29(1):436–460.