An unbiased, open-source, algebraic **method** *to* **rank** college football teams *and* **predict** future college football game outcomes using statistics

By Neville Aga @playoffPredict
December 2022

Abstract:
A simple method to rate college football teams within a season is proposed and analytically discussed. The method uses only one input (final scores of games within a given season) and produces a "football truth", an accurate, unbiased ranking of all 131 FBS (football bowl subdivision) teams. Results are then compared to betting markets and human polls as evidence of the correctness of the model.

Table of Contents

1.0 Introduction

The goal of this paper is to introduce a computational, bias-free, and open-source method to:

#1 Accurately Rate and Rank each FBS team, and

#2 Predict what the playoff committee will list as their top teams each week

Additionally, after ratings and rankings are computed by this method, those results can be used to

#3 Predict winners in future college football matchups.

The largest overall fundamental premise to this method is the mathematical simplicity. If successful, I will have persuaded the reader that the method is mathematically elegant[1] – pleasingly ingenious and simple. To that end the playoffPredictor.com computer rating method uses the single input of final scores in games for the current season to create rankings. Let's state that again: the premise of this method is the underlying math is so correct and intuitive that starting from nothing each season an accurate rating, ranking, and prediction model can be built with the simple inputs of final scores in games played during that season.

The method aims to use no inputs that are subjective such as offensive points scored, total yards gained, replacement value of a backup quarterback, etc. The reason that these metrics are subjective and bad choices for an elegant mathematical method is because the method maker must determine what constitutes "good" for those categories. Of course, a 500-yard offensive performance is accepted to be considered good, but how good and how does it compare to a turnover margin of -2? Combining many inputs leads to noise, which leads to subjectivity in any computer-based formula. To do a complete mathematical analysis using such inputs there would need to be a mechanism to relate categories to each other – for example zero punts in a game is worth two touchdowns. These are valuable efforts and lead to better predictive models, but still subjective in how the model maker combines these inputs to arrive at final ratings. In practice from other model-makers, weights are assigned by back-testing the full output results to historical data. As the SEC often makes its member's say, "past performance is no guarantee of future results" (no, not that SEC, I of course mean the Securities and Exchange Commission).

SECTION 1 – COMPUTER MODEL TO RATE AND RANK TEAMS

[1] Elegant, of a scientific theory or solution to a problem, as in "the grand unified theory is compact and elegant in mathematical terms"

One input that is not subjective is wins and losses. Football is a team sport and at the end of the day the only thing that should be necessary and sufficient to rank all teams is wins and losses against the scheduled played. A winning outcome is the goal of the team and the only metric that need be considered to make a good model on ranking teams.

So how do we map this to college football? The best model is the one proposed by Pierre-Simon Laplace in 1814. Yes, that Laplace you studied in college math of Laplace transformation fame. Of course, Laplace did not apply his method to college football since he preceded it by a hundred years[2]. Instead, Laplace sought to answer the question “Will the sun rise tomorrow?” Which turns out to be a good fit to answer the question “Will my football team win next week?”.

A football team winning or losing a single game is a binomial probability. In each trial (football game) there is only success (W) or failure (L). Will the sun rise tomorrow can also be modeled as a binomial – success (it will rise), or failure (it will not rise). Now, having no further insight except that the trial must end in success or failure you can model this probability by

$$P(sun\ will\ rise\ tomorrow) = \frac{d+1}{d+f+2} \tag{1}$$

where d represents the number of times the sun has risen in the past, and f represents the mornings where the sun did not rise. Clearly the sun has never (yet) failed to rise, so equation 1 simplifies to

$$P(sun\ will\ rise\ tomorrow) = \frac{d+1}{d+2}$$

So, on day 1 when God created Adam, and subsequently Adam wondered if the sun would rise tomorrow, he would have computed the probability as ½ - having no prior data and no knowledge of the workings of gravity, etc. – it’s a 50/50 shot. On day 2 after a successful first sunrise the odds for day #3 improve to 2/3, and by now with 3,000,000 days where humans have documented the sun did indeed rise yesterday, the odds for tomorrow improve substantially to

$$\frac{3{,}000{,}000}{3{,}000{,}001} = 0.999999667 \approx 1$$

Now of course Laplace had insight to say that if you understand the mechanics of gravity and planetary motion you can make a much better guess as to the true probability. We will leave the detailed models to the other model-makers. Our goal is can we be simple in our inputs and arrive at some type of “football truth”?

[2] And he was French, and the French have no love for American college football

Of special note, the reader should recognize the formula for the sun rising is not $\frac{d}{d}$, it is $\frac{d+1}{d+2}$. This is a critical detail. Without that 1 in the numerator and 2 in the denominator what you will find out is that probability models do not converge. Laplace treated the "why" mathematically, and Wes Colley in his seminal paper also explains in good detail why the ½ is necessary for a start. It is not in the scope of this paper to prove the necessity of the ½, please see Wes Colley's 2002 paper if the reader needs proof of that fact.

At this point we switch from probability notation [P(X)] to rating notation [r_X]. We are transforming the probability into a rating, and the rating is only valid in **pairwise operations**. That is, unlike probabilities which stand on their own [P(heads)=0.5], a rating for a team is only valid when comparing against the rating of another team [for instance, $r_{Auburn} = 0.7$ & $r_{Oklahoma} = 0.6$]. On its own $r_{Auburn} = 0.7$ means nothing, it must be compared to the other ratings in the system. Indeed, when we make our corrections for ratings for strength-of-schedule and margin-of-victory our ratings can take a value greater than 1 or less than 0. Probability notation will not suffice.

2.0 - Basic 2 team analysis and strength of schedule

To map to football, simply use the idea of the sun successfully rising as a win (football number of wins is analogue to sun successfully rising) and sun not rising as a loss (football number of losses is analogue to sun not rising for one given morning). In our notation d becomes n_w and f becomes n_l . The ratings for team A become

$$r_A = \frac{n_{w,A} + 1}{n_{w,A} + n_{l,A} + 2} \tag{2}$$

So, we start with simple wins and losses. That gives us a winning percentage and a way to rank teams. The first obvious limitation is simple winning percentage does not consider strength of schedule. If we consider 12-0 Alabama and trying to rank that against 12-0 Cincinnati, for both teams we will compute their rating as $\frac{12+1}{12+2} = \frac{13}{14} \approx 0.929$

We need a way to understand that Alabama's 12 wins are superior to Cincinnati's 12 wins. Strength of schedule is the first step to get there.

2.1 Adding in Strength of Schedule (the Colley method)

Note that number of wins can be rearranged:

$$n_w = \frac{(n_w - n_l)}{2} + \frac{n_{tot}}{2}$$

(Which the reader can check). Recognize the second term may be written as

$$\sum_{n_{tot}} \frac{1}{2}$$

allows one to identify the sum as that of the ratings of a team's opponents if those opponents are all random (r = 1/2) teams. Instead, then, of using r = 1/2 for all opponents, we now use their actual ratings, which gives an obvious correction to n_w. (now using the term $n_{w,i}$ to mean team "i" is under consideration)

$$n_{w,i}^{eff} = \frac{(n_{w,i} - n_{l,i})}{2} + \sum_{j=1}^{n_{tot,i}} r_j^i \tag{3}$$

where r_j^i is the rating of the j^{th} opponent of team i. The second term (the summation) in equation (3) is the adjustment for strength of schedule[3].

Axiom #1: The sum of rating of random teams can be replaced with the sum of the rating of teams played since $\bar{r} = 0.5$

The resulting rating formula with SoS becomes

$$r_A = \frac{1 + n_{w,A}^{eff}}{2 + n_{w,A} + n_{l,A}} \tag{4}$$

The goal is to simultaneously solve all the r_j^i's which are inputs to the r_i's. What we end up with is a system of 131 equations and 131 unknowns (131 being the number of football teams in a given FBS season). That becomes an algebraically solvable system. Thanks to modern computing power, solving a system of equations (that has a bound solution) is a simple task.

Notice the denominator in equation 4 is simply the number of games A plays plus 2. Unlike the numerator there is no concept of effective wins (SoS). That allows a straightforward, uncomplicated matrix solution to the equations.

Let's apply this math to a simple two-team league. In our simple world, team A plays and beats team B. The system of equations describing this league would simplify to:

---

[3] This equation and the surrounding paragraphs are copied directly from Wes Colley's 2002 paper. The credit for understanding ½ random should be replaced with actual ratings is the first and most important mathematical elegance of this method. In my opinion, that insight is like an $E = mc^2$ type of insight. Simple, mathematically correct, elegant.

$$r_A = \frac{1+n_{w,A}^{eff}}{2+n_{tot,A}} \qquad r_B = \frac{1+n_{w,B}^{eff}}{2+n_{tot,B}}$$

where

$$n_{w,A}^{eff} = \frac{(n_{w,A} - n_{l,A})}{2} + r_B \qquad n_{w,b}^{eff} = \frac{(n_{w,B} - n_{l,B})}{2} + r_A$$

Substituting in the equality for n_w^{eff},

$$r_A = \frac{1+\frac{(n_{w,A} - n_{l,A})}{2}+r_B}{2+n_{tot,A}} \qquad r_B = \frac{1+\frac{(n_{w,B} - n_{l,B})}{2}+r_A}{2+n_{tot,B}}$$

$$r_A = \frac{1+\frac{(1-0)}{2}+r_B}{2+1} \qquad r_B = \frac{1+\frac{(0-1)}{2}+r_A}{2+1}$$

$$r_A = \frac{1.5+r_B}{3} \qquad r_B = \frac{0.5+r_A}{3} \tag{5}$$

Two equations and two unknowns. The reader can check that $r_A = 0.625$ and $r_B = 0.375$ satisfy these equations exactly.

Arranging the first writing of equation (5) for team A differently,

$$(2 + n_{tot,A})\, r_A - r_B = 1 + \frac{(n_{w,A} - n_{l,A})}{2} \tag{6}$$

Extending to 131 FBS teams playing 800-900 games on average it is convenient arrange equation (6) in summation form:

$$\left(2+n_{tot,i}\right) r_i - \sum_{j=1}^{n_{tot,i}} r_j^i = 1 + \frac{\left(n_{w,i} - n_{l,i}\right)}{2}$$

If desired, isolate the rating:

$$r_i = \frac{1 + \frac{\left(n_{w,i} - n_{l,i}\right)}{2} + \sum_{j=1}^{n_{tot,i}} r_j^i}{\left(2+n_{tot,i}\right)}$$

Switch to matrix form by rewriting equation (5) as follows,

$$C\vec{r} = \vec{b}, \tag{7}$$

where $\vec{r}$ is a column-vector of all the ratings r_i, and $\vec{b}$, is a column-vector of the right-hand-side of equation (6):

$$b_i = 1 + \frac{(n_{w,i} - n_{l,i})}{2}$$

The i^{th} row of matrix C has as its i^{th} entry $2 + n_{tot,i}$, and a negative entry of the number of games played against each opponent j. In other words,

$$\begin{aligned} c_{ii} &= 2 + n_{tot,i} \\ c_{ij} &= -n_{j,i} \end{aligned}$$

Solving the preceding equations is the method for the Colley Matrix rating of the teams.

2.2 Adding in Margin of Victory (the playoffPredictor.com method)

The next logical question is if the Colley Matrix method is sufficient. Ideally it would be, but in practice it is not quite, in my estimation. Or at least I feel we can and need to do better. The method starts with all teams at $r = 0.500$ and only converges at very good ratings by the end of the season. If we want to predict future games midseason it does us no good to know that the ratings will be right at the end of the season! Furthermore, Colley defined an error statistic η, that by the end of the season with his method got to an error of about ~1.25 by the end of the season, and is about ~1.6 midseason. Alas, if a college football season was ~20-25 games long, then the Colley method alone would converge with press polls, or true ratings beautifully. However, we need a way to make these ratings converge with press ratings by early or midseason to test the predictive powers of the model. Enter margin of victory.

Our next task is to introduce a set of coupled variables; one that extends the definition of an effective win ($n_{w,A}^{eff}$) using margin-of-victory, and another variable that relates this margin-of-victory to the already defined win/loss matrix.

2.2.1 Margin of Victory method – win margin perspective

At this point we introduce a new variable m and an associated weighting α. We extend the equations by adding an extra term to the rating:

$$r_i = \frac{1 + \left(n_i^{w,eff} + \alpha\, m_{tot,i}^{eff}\right)}{2 + \left(n_i - \alpha\, m_{tot,i}\right)} \tag{8}$$

$n_i^{w,eff}$ is defined as it was in eq 3 (slight notation difference noted). We define $m_{tot,i}^{eff}$ as:

$$m_{tot,i}^{eff} = \sum_{j=1}^{n_{tot,i}} r_j^i \cdot m_{ij}$$

n_i is the total number of games team i has played (can also be written as n_i^{tot} or $n_{tot,i}$).

$m_{tot,i}$ is the total number of margin team i has accumulated in all games played (can also be written as m_i^{tot}). Margin for each game is a number between $[-1{,}1]$ so the sum of all margins for a team i will be between $[-n_i, +n_i]$.

$$m_{tot,i} = \sum_{j=1}^{n_{tot,i}} m_{ij}$$

The new term m_{ij} is defined as the margin of victory in the game played between teams i and j. m_{ij} is a positive number for a win by team i and an equal but negative number for m_{ji} as a loss[4]. Again, m_{ij} can have any value between $[-1, +1]$

α is a scaling factor that relates margin-of-victory to wins and losses (which we will solve for later).

What we are doing is modifying the number of effective wins and number of total wins by some m margin. Since the m margin is related to the n games, we do not need to add an entire $(1 + W)/(2 + T)$ form again, we can add simply extend the $n_{w,A}^{eff}/n_{tot,A}$ form into a $\left(n_{w,A}^{eff} + m_{w,A}^{eff}\right)/\left(n_{tot,A} - m_{tot,A}\right)$. Notice the extension is additive in the numerator while negative in the denominator.

When thinking about it from the point of view of a "hidden" maker on a unit width craps table, we are now being told not only if our cast die landed to the left of the marker, but we are also told *how far left* it was cast. We still our not told where our hidden marker (true team

[4] The resulting matrix is not symmetrical anymore since $m_{ji} \neq m_{ji}$, but it is still of full rank and solvable

rating) is, but we are told "you landed left (won the game), but very close to the marker (m_{ij} is very small and positive)", or "you landed left (still won the game), far away from the marker (m_{ij} is very large and positive)". If you lose the game (land to the right of the marker) then m_{ij} is a negative number, either small or large depending on how badly you lost. This is the conceptual value of m, our margin of victory. When defined like this $m_{ij} = -m_{ji}$, that is team i's win is of equal magnitude and opposite sign to team j's loss.

Note inside m_i^{eff} that absolute score values are not considered – that is there is no m_i^{tot} for the team. The only margins are multiplied by the team played – indeed blowing out Duke by 23 points is very different from blowing out Ohio State by 23 points.

The $\sum_{j=1}^{n_{tot,i}} r_j^i \cdot m_{ij}$ component is defined from the perspective of wins, so if a team wins all their games this will be a large positive number. If they win half and lost half, this will be about 0, and if they lose all their games this will be a large negative number. Drilling down further, note that if a team A beats a top team by the same margin they lose to a bottom team that the net result will be positive for team A. That is why we define margin-of-victory this way as a ***win-margin perspective***.

Recalling eq 3 for $n_{w,i}^{eff}$ and expanding out the full form of eq 7 we get:

$$[2 + n_{tot,i} - \alpha\, m_{tot,i}]\, r_i = 1 + \left[\frac{(n_{w,i} - n_{l,i})}{2}\right] + \left[\sum_{j=1}^{n_{tot,i}} r_j^i\right] + \alpha\left[\sum_{j=1}^{n_{tot,i}} r_j^i \cdot m_{ij}\right]$$

Collecting our r terms:

$$[2 + n_{tot,i} - \alpha\, m_{tot,i}]\, r_i - \sum_{j=1}^{n_{tot,i}} [1 + \alpha\, m_{ij}]\, r_j^i = 1 + \left[\frac{(n_{w,i} - n_{l,i})}{2}\right]$$

Meaning that the full final rating for team A is given as:

$$r_A = \frac{1 + \frac{(n_{w,A} - n_{l,A})}{2} + \sum_{j=1}^{n_{tot,A}} [1 + \alpha\, m_{Aj}]\, r_j^A}{(2 + n_{tot,A} - \alpha\, m_{tot,A})}$$

Solving that set of equations for all 132 teams A-Z is the playoffPredictor.com win-perspective computer method for rating the teams.

A value of $m_{AB} = 0$ means that team A and B finished with no margin of victory (like a 1-point win), which effectively makes the playoffPredictor.com computer method simplify into the Colley method. A value of 1 for m_{AB} means that team A beat team B with full margin-of-victory achieved. Think team A beating team B in a 100-point blowout. We will quantify this thinking now:

Let's say team A beats team B by the final score of 101-0. We want to give them more credit for this win than just a single standard win. In our model 101 points will equate to a full margin-of-victory credit ($m = 1$), which will make the math simpler. For argument's sake we can assume a linear relationship for m such that a 31-point victory would have resulted in $m = 0.3$. We do not know the correct value of α yet, but we will use 0.5. The rating of team A becomes:

$$[2 + 1 - \alpha\,(1)]\, r_A \;-\; [1 + \alpha\,(1)]\, r_B \;=\; 1 + \frac{(1)}{2}$$

$$[3 - \alpha\,]\, r_A \;-\; [1 + \alpha\,]\, r_B \;=\; 1.5$$

$$r_A \;=\; \frac{\left[1.5 + \frac{\alpha}{2}\right] \;+\; [1 - \alpha\,]\; r_B}{[3 - \alpha\,]}$$

The solution for these pair of equations for $\alpha = 0.5$ is

$$r_A \;=\; 0.75; \quad r_B \;=\; 0.25$$

In matrix form, the equations become:

$$\vec{r} = A^{-1}\vec{d}, \tag{9}$$

Note the difference to equation (7). We are trying to solve for ratings, so the ratings variable is isolated on the left-hand side. We use the notation A^{-1} to denote the inverse of the A matrix, which is the C matrix extended for margin-of-victory. $\vec{r}$ is still the column-vector of all the team's ratings r_i, and $\vec{d}$, is a column-vector of:

$$d_i = 1 + \frac{\left(n_{w,i} - n_{l,i}\right)}{2} \tag{10}$$

The i^{th} row of matrix A has as its i^{th} entry $2 + n_{tot,i} - \alpha\,(m_{tot,i})$, and an entry of [-1 + the MoV scaling factor (α) * the MoV value m] for each game played against each opponent j. In other words,

$$\begin{aligned} a_{ii} \;&=\; 2 \;+\; n_{tot,i} - \alpha\,(m_{tot,i}) \\ a_{ij} \;&=\; -\,n_{i,j} \;+\; \alpha\,(m_{ij}) \end{aligned} \tag{11}$$

The matrix A is the **playoffPredictor.com Matrix**. It extends the Colley Matrix with information on margin of victory. *Solving equations (9)–(11) is the playoffPredictor.com win-perspective computer method for the rating of the teams in the matrix domain.*

In our 2-team scenario the non-MoV matrix is

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} .625 \\ .325 \end{bmatrix}$$

Adding in α and m as described results in:

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} 3 - \alpha \cdot m_{AB} & -1 + \alpha \cdot m_{BA} \\ -1 + \alpha \cdot m_{AB} & 3 + \alpha \cdot m_{BA} \end{bmatrix}^{-1} \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} .625 + \alpha \cdot m_{AB}/3 \\ .375 + \alpha \cdot m_{BA}/3 \end{bmatrix}$$

For $\alpha = 0.5$ and $m_{AB} = 1$ / $m_{BA} = -1$ (full MoV victory) the matrix becomes:

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} 3 - 0.5 & -1 - 0.5 \\ -1 + 0.5 & 3 + 0.5 \end{bmatrix}^{-1} \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$$

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} 2.5 & -1.5 \\ -0.5 & 3.5 \end{bmatrix}^{-1} \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} .75 \\ .25 \end{bmatrix}$$

r_A has changed from 0.625 in a non MoV context to 0.75 in a MoV context.

Notice when α and m added in this way we preserve the nature of the matrix, namely

$\bar{r} = 0.5$, (the average rating of all teams is 0.5)
$\sum d = count_{teams}$,
$\sum A^{-1} = 0.5 * count_{teams}$, and
$\sum A = 2 * count_{teams}$

If those criteria are met the matrix remains stable and solvable.

There is no elegant truth to the correct value for m for any given game. The model-maker can choose to scale it linearly such that a 2-point victory[5] is worth $m = 0.02$ and a 100-point victory is worth $m = 1.0$, or the model-maker can use arctan functions to defeat runaway scores, logistic functions, or a discrete mapping. There is solid logic behind a stepwise (nonlinear) increase between a margin of victory of 8 points and 9 points – the former can be tied on one score at the end of the game, where the later needs a score, an onside kick, and another score.

Ideally, m should be based on how long it takes for team A to defeat team B in the time domain, not points scored in the scoreboard. As a practical matter it is hard to come up with that value, and score is well correlated to "time of victory", a measure of when team A is beating team B so

[5] A 1-point victory is expressly defined as $m = 0$, so the equation $\vec{r} = A^{-1}\vec{d}$, expressly simplifies to $\boldsymbol{\vec{r}} = \boldsymbol{C^{-1}\vec{b}}$

bad, team B stops trying and just runs the ball and punts. The ideal mapping would see m would scale linearly with the time it takes team A to put away team B.

Since we don't know the truth behind different margin-of-victory outcomes, we just use a simple linear scaling:

Effective MoV mapping

MoV	Effective MoV
1	0
2	0.01
3	0.02
4	0.03
5	0.04
6	0.05
7	0.06
8	0.07
9	0.08
10	0.09
11	0.1
12	0.11
13	0.12
14	0.13
15	0.14
16	0.15
17	0.16

We start with a 1-point victory resulting in $m = 0$, and we just linearly scale up from there by .01 per point. So, in the game between team i and team j:

$$m_{ij} = \frac{total\ points\ scored\ by\ team\ i\ - total\ points\ scored\ by\ team\ j - 1}{100}$$

if team i wins, and

$$m_{ij} = \frac{total\ points\ scored\ by\ team\ i\ - total\ points\ scored\ by\ team\ j + 1}{100}$$

if team i loses.

This perspective (the win perspective) rewards teams for large margin of victory over good teams. The rating for a give team includes the term $\sum_{j=1}^{n_{tot,A}} \left[1 + \alpha\, m_{Aj}\right] r_j^A$. If margin is held steady, a larger boost comes from defeating a team with a higher r_j. This is in line with common sense. However, what happens if a team has a bad loss – a loss to a bad team by a large (negative) m? They will be punished less for an m loss to a bad team, and punished more for an m loss to a good team. In fact, comparing $r_{jbad} = 0.3$ to $r_{jgood} = 0.9$ the MoV punishment is 3 times as great for the loss to the good team over the bad team. That is counter to common sense.

2.2.2 Margin of Victory alternate candidates

Keeping the same basic form of Strength-of-Schedule, we can add Margin-of-Victory from a loss perspective instead of a win perspective. This method will seek out losses and make bad teams have ratings around 1.0 and great teams have ratings around 0.0. Success in this trial is a loss.

$$\acute{r}_i = \frac{1 + \left(n_{l,i}^{eff} + \alpha\, m_{l,i}^{eff}\right)}{2 + \left(n_{tot,i} + \alpha\, m_{tot,i}\right)}$$

Here we use $\acute{r}_i$ instead of r_i. Since we are solving for losses, bad teams will have an $\acute{r}_i$ close to 1.0 and good teams will have an $\acute{r}_i$ close to 0.0. We don't want to get these confused with r_i because we will transform these $\acute{r}_i$ into r_i at the end of our method to keep similar senses of good and bad. The denominator is changed to a + since $m_{tot,i}$ is still negative for bad teams.

$$n_{l,i}^{eff} = \frac{\left(n_{l,i} - n_{w,i}\right)}{2} + \sum_{j=1}^{n_{tot,i}} \acute{r}_j^{i}$$

Keeping the convention that m_{ij} is still a negative number for a loss of i to j, we now have a new definition of $m_{l,i}^{eff}$ as:

$$m_{l,i}^{eff} = -\sum_{j=1}^{n_{tot,i}} \acute{r}_j^{i} \cdot m_{ij}$$

Expanding out the full form of $n_{w,i}^{eff}$ and $m_{l,i}^{eff}$ we get:

$$\left[2 + n_{tot,i} + \alpha\, m_{tot,i}\right] \acute{r}_i = 1 + \left[\frac{\left(n_{l,i} - n_{w,i}\right)}{2}\right] + \left[\sum_{j=1}^{n_{tot,i}} \acute{r}_j^{i}\right] - \alpha \left[\sum_{j=1}^{n_{tot,i}} \acute{r}_j^{i} \cdot m_{ij}\right]$$

$$\left[2 + n_{tot,i} + \alpha\, m_{tot,i}\right] \acute{r}_i - \sum_{j=1}^{n_{tot,i}} \left(1 - \alpha\, m_{ij}\right) \acute{r}_j^{i} = 1 + \left[\frac{\left(n_{l,i} - n_{w,i}\right)}{2}\right]$$

What this will accomplish is drawing out the blowout losses to bad teams. The final rating now has the term -$\sum_{j=1}^{n_{tot,i}} \left(-\alpha\, m_{ij}\right) \acute{r}_j^{i}$ (1 intentionally omitted, they have to play each other first). Just talking about what MoV does) in the numerator. If team i loses by 101 points to team j with a rating of $\acute{r}_j = 0.8$ (meaning $r_j = 0.2$), and using $\alpha = 0.5$ the loss-rating for team i will decrease by $(-.5(-1)) \cdot 0.8$. We are getting a much larger final weighting when using 0.8 instead of 0.2

Finally, we use a transformation of $r_i = 1 - \acute{r}_i$ to get final ratings with this method.

Another method for computer ratings can be formulated by realizing that an average team has a rating of 0.5, therefore 1-the rating also will be 0.5 Again, this perspective can be useful to draw out information on weaker teams. We can redefine effective wins as

$$n_{l,i}^{eff} = \frac{(n_{l,i} - n_{w,i})}{2} + \sum_{j=1}^{n_{tot,i}} (1 - \acute{r}_j^i)$$

and $m_{l,i}^{eff}$ as:

$$m_{l,i}^{eff} = \sum_{j=1}^{n_{tot,i}} (1 - \acute{r}_j^i) \cdot m_{ij}$$

Axiom #2: The sum of rating of random teams can also be replaced with the sum 1 minus the rating of teams played since 1-$\bar{\acute{r}} = 0.5$

Expanding out the full form of $n_{w,i}^{eff}$ and $m_{l,i}^{eff}$ we get:

$$[2 + n_{tot,i} + \alpha\, m_{tot,i}]\, \acute{r}_i + \sum_{j=1}^{n_{tot,i}} (1 + \alpha\, m_{ij})\, \acute{r}_j^i = 1 + \left[\frac{(n_{l,i} - n_{w,i})}{2}\right] + n_{tot,i} + \alpha\, m_{tot,i}$$

Generally, I am unhappy with all permutations of (1-r). However, when α and m added in this way we have a different nature of the matrix, namely $\bar{\acute{r}} \neq 0.5$, unless all teams have played the same number of games. College football does not have teams playing equal games with conference championships, bowl games, and cancelled games. The method may be good enough, but it is not perfect, and hence not considered for an elegant method.

Finally, there is a method to go totally nonlinear. This method can only be solved iteratively and not directly from matrix form:

$$r_i = \frac{1 + \left(n_i^{w,eff} + \alpha\, m_i^{w,eff}\right)}{2 + \left(n_i^{w,eff} + \alpha\, m_i^{w,eff}\right) + \left(n_i^{l,eff} + \alpha\, m_i^{l,eff}\right)}$$

$$with\ the\ following\ definitions:$$

$$n_{w,i}^{eff} = \frac{(n_{w,i} - n_{l,i})}{2} + \textstyle\sum_{j=1}^{n_{w,i}} r_j^i \qquad m_i^{w,eff} = \textstyle\sum_{j=1}^{n_{w,i}} r_j^i \cdot m_{ij}$$

$$n_{l,i}^{eff} = \frac{(n_{l,i} - n_{w,i})}{2} + \textstyle\sum_{j=1}^{n_{tot,i}} (1 - r_j^i) \qquad m_{l,i}^{eff} = \textstyle\sum_{j=1}^{n_{tot,i}} (1 - r_j^i) \cdot m_{ij}$$

So ultimately which model is correct for adding in margin-of-victory? There is no absolute answer. Generally, the win-model tracks better to the polls, probably because there is a lot of studying among top teams, much more than comparative studying between UConn and New Mexico State. But if we are looking for football truth, alas none can be found.

At this time the playoffPredictor.com computer model uses the win-margin perspective only.

2.3 Theoretical value of α

Is there a way to deduce the correct value of α theoretically? Remember α relates MoV to winning a game. Stated another way, there is proportionality between a game win and a high MoV [$p(win) \propto +MoV$]. Alpha is a constant that makes this proportion into an equation [$p(win) = f(\alpha, (+MoV))$]. How is it possible to theoretically predict the margin of victory for a winner? There is a linkage through rematches. Note what happens in our 2-team league with rematches. If A plays B two times and wins both times we get

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} .667 \\ .333 \end{bmatrix}$$

That is r_A has increased from 0.625 to 0.667. In general, for R successive rematches, all won by A, the rating column-vector without MoV goes to:

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} 3+R & -1-R \\ -1-R & 3+R \end{bmatrix}^{-1} \begin{bmatrix} 1.5+0.5R \\ 0.5-0.5R \end{bmatrix}$$

For many repeated rematches and wins by A we approach

$$\begin{bmatrix} r_A \\ r_B \end{bmatrix} = \begin{bmatrix} \infty & -\infty \\ -\infty & \infty \end{bmatrix}^{-1} \begin{bmatrix} {}^{1}/_{2}\ \infty \\ -{}^{1}/_{2}\ \infty \end{bmatrix} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$$

	1 point victory (α = 0, no MoV in these wins)						
	# matches played						
teams	1	2	3	10	30	100	∞
A	0.625	0.667	0.688	0.727	0.742	0.747	0.750
B	0.375	0.333	0.313	0.272	0.258	0.252	0.250

Figure 1: Progression of team ratings of A&B with repeated A vs B rematches with A winning every time

Returning to our formula with MoV, let's see how different values of α (holding m steady at 1) affect $\vec{r}$:

1 point victory (α = 0, no MoV in these wins)

		# matches played
	teams	1
	A	0.625
	B	0.375

100 point victory (m =1), α = -.1

		# matches played
	teams	1
	A	0.645
	B	0.355

100 point victory (m= 1), α = -.25

		# matches played
	teams	1
	A	0.679
	B	0.321

100 point victory (m = 1), α -.5

	teams	# matches played
		1
	A	0.750
	B	0.250

Note that *for an* $\alpha = 0.5$ *we get the ratings with our full MoV after 1 game equal to the ratings in a non-MoV environment after infinite rematches*. A full margin-of-victory win is equal to saying that team A would beat team B 100% of the time. Is this reasonable? From a theoretical standpoint, yes! We can set MoV however we want to, but if team A can beat team B by 100 points or something absurd like that, we can say that's equivalent to team B never beating team A, even if you gave them a thousand tries. Their probability spaces do not overlap anymore.

Axiom #3: The MoV to win scaling constant, alpha, is defined as $\alpha = 0.5$

It is critical to appreciate that we have arrived at α without any historical back testing whatsoever. To this point we are all in the realm of theory. In practice, $\alpha = 0.5$ is a good fit to the playoff committee / AP poll, but actually $\alpha = 0.3$ is a better fit to the data. Regardless we will use $\alpha = 0.5$ because our stated goal is to see if there is an algebraic way to arrive at "football truth"[6].

[6] Or it could simply be that the pollsters are wrong, and the computer is right. In general, as α goes from 0 to ~0.1 there is an incredible increase in the accuracy of η – improvements around 20%-35%. Generally, η sees maximum improvement around the range of $0.2 < \eta < 0.333$, which are as high as high as 50%. At $\alpha = 0.5$ η sees improvement around 20-30%

2.4 Common Sense calculations that should be satisfied

1) If team A beats team B by 2 points, the resulting r_A must be greater than if team A had beat team B by only 1 point.

2) Assume team A has 12 wins and no losses. Two of those wins are by a score of 7-0 over a top team and a bottom team. If the win over the bottom team is moved to 8-0 holding the top win steady at 7-0, r_A should get a boost, but that boost must be less than the boost that it gets if its win over the bottom team is returned to 7-0 and the win over the top team is moved to 8-0.

3) Assume team A has 10 wins 2 losses. The two losses are to a top team by a score of 0-7 and to a bottom team also by the score of 0-7. If the loss to the top team is moved to 0-8 holding the bottom loss steady at 0-7, r_A should get a decline, but less than the decline when the loss to the top team is returned to 0-7 and the loss to the bottom team is moved to 0-8.

Both the win perspective and the loss perspective accomplish #1. The win perspective accomplishes #2, and the first proposed loss-perspective accomplishes #3. *If desired, we could simply average the win-perspective and loss-perspective rating to arrive at the final ratings for each team.* That may satisfy all common-sense calculations.

3.0 moving the model to 3 teams

Consider a 3-team league, teams A-C. Three games are played, A beats both B and C by 1 point and B beats C, also by one point. Our equations become:

$$[2 + n_{tot,A} - \alpha(m_{tot,A})]\, r_A = 1 + \left[\frac{(n_{w,A} - n_{l,A})}{2} + r_B + r_C\right] + \alpha[\, -r_B m_{AB} - r_C m_{AC}]$$

$$[2 + n_{tot,B} - \alpha(m_{tot,B})]\, r_B = 1 + \left[\frac{(n_{w,B} - n_{l,B})}{2} + r_A + r_C\right] + \alpha[\, -r_A m_{BA} - r_C m_{BC}]$$

$$[2 + n_{tot,C} - \alpha(m_{tot,C})]\, r_C = 1 + \left[\frac{(n_{w,C} - n_{l,A})}{2} + r_A + r_B\right] + \alpha[\, -r_A m_{CA} - r_B m_{CA}]$$

$$[2 + 2 - 0.5(0)]\, r_A = 1 + \left[\frac{(2-0)}{2} + r_B + r_C\right] + 0.5[\, -r_B \cdot 0 - r_C \cdot 0]$$

$$[2 + 2 - 0.5(0)]\, r_B = 1 + \left[\frac{(1-1)}{2} + r_A + r_C\right] + 0.5[\, -r_A \cdot 0 - r_C \cdot 0]$$

$$[2 + 2 - 0.5(0)]\, r_C = 1 + \left[\frac{(0-2)}{2} + r_A + r_B\right] + 0.5[-r_A \cdot 0 - r_B \cdot 0]$$

$$[4]\, r_A = [2 + r_B + r_C] + [0]$$
$$[4]\, r_B = [1 + r_A + r_C] + [0]$$
$$[4]\, r_C = [0 + r_A + r_B] + [0]$$

The values $r_A = 0.7$, $r_B = 0.5$, $r_C = 0.3$ solve these equations. This is intuitive. A is separated from B by the same distance as B is separated from C.

Now, consider what happens if A beats B by 21, A beats C by 41, and B beats C by 21. In our model a 20-point victory gives $m = 0.5$ and a 40-point victory gives $m = 1$:

$$[2 + n_{tot,A} - \alpha(m_{tot,A})]\, r_A = 1 + \left[\frac{(n_{w,A} - n_{l,A})}{2} + r_B + r_C\right] + \alpha[\,-r_B m_{AB} - r_C m_{AC}]$$

$$[2 + n_{tot,B} - \alpha(m_{tot,B})]\, r_B = 1 + \left[\frac{(n_{w,B} - n_{l,B})}{2} + r_A + r_C\right] + \alpha[\,-r_A m_{BA} - r_C m_{BC}]$$

$$[2 + n_{tot,C} - \alpha(m_{tot,C})]\, r_C = 1 + \left[\frac{(n_{w,C} - n_{l,A})}{2} + r_A + r_B\right] + \alpha[\,-r_A m_{CA} - r_B m_{CA}]$$

$$[2 + 2 - 0.5(.6 - 0)]\, r_A = 1 + \left[\frac{(2-0)}{2} + r_B + r_C\right] + 0.5[\,-r_B \cdot 0.2 - r_C \cdot 0.4]$$

$$[2 + 2 - 0.5(.2 - .2)]\, r_B = 1 + \left[\frac{(1-1)}{2} + r_A + r_C\right] + 0.5[\,-r_A \cdot 0.2 - r_C \cdot 0.2]$$

$$[2 + 2 - 0.5(0 - .6)]\, r_C = 1 + \left[\frac{(0-2)}{2} + r_A + r_B\right] + 0.5[\,-r_A \cdot 0.4 - r_B \cdot 0.2]$$

$$\begin{aligned}
[3.7]\, r_A &= [2 + r_B + r_C] + [-0.1 r_B - 0.5 r_C] \\
[4]\, r_B &= [1 + r_A + r_C] + [-0.25 r_A - 0.25 r_C] \\
[4.3]\, r_C &= [0 + r_A + r_B] + [-0.5 r_A - 0.25 r_B]
\end{aligned}$$

$$\begin{aligned}
3.7\, r_A - 1.1 r_B - 0.5 r_C &= 2 \\
4\, r_B - 0.75 r_A - 0.75 r_C &= 1 \\
4.3\, r_C - 0.75 r_B - 0.5 r_A &= 0
\end{aligned}$$

The values $r_A = 0.766$, $r_B = 0.490$, $r_C = 0.245$ solve these equations.

Look at what we are doing as m increases from the point of view of team A. Because effective margin is subtracting the from the coefficient for B & C, as margin goes up, more of the rating from B or C has to be transferred in order to satisfy the equations. In the above example since A beat C twice as bad as B, the coefficient in front of r_C is twice as far from -1 as the coefficient in front of r_B for equation 1 (that is -1.2 compared to -1.1).

Standard deviations

Consider the 3-team league where A beats both B&C and B beats C. Clearly A will be on top every time, but by what margin? In the case of all 1-point victories we saw $r_A = 0.7, r_B = 0.5, r_C = 0.3$. In the case of the 41-point victory A over C and 21-point victories A-B & B-C, we saw $r_A = 0.867, r_B = 0.5, r_C = 0.133$. If we look at these two scenarios from the view of standard deviations, we see they are both the same result: A is +1.22 standard deviations over B, and C is -1.22 standard deviations under B. Mathematically there is no difference in relative goodness between these teams if the only information we have is 20-point victories or 1-point victories, if they all scale together. However, the model does produce interesting results if there are inconsistencies in the victories. For example, if A barely beats B, and A barely beats C, you expect B to basically be tied with C and the game between B and C close. But what if it is not close, what if B destroys C. Then, to satisfy the equations, you arrive at $r_A = 0.7, r_B = 0.588$, $r_C = 0.213$. Notice that A does not get the 'credit' for the B destruction of C. If all 3 games were 'barely beat', then the model produces the aforementioned .7/.5/.3. But because B>>C, B's rating is increased by .088 at the expense of C. In a Z score context, A is +0.96, where B is +.42, and C is -1.38. See the chart below for all the permutations of 3 teams considering margin of victory of "barely beats" ($m = 0$), "beats" ($m = 0.5$), and "destroys" ($m = 1$).

									rA	rB	rC	σA	σB	σC
A	destroys	B	B	barely beats	C	A	barely beats	C	0.788	0.413	0.300	+1.38	-0.42	-0.96
A	destroys	B	B	barely beats	C	A	beats	C	0.850	0.406	0.245	+1.37	-0.37	-1.00
A	destroys	B	B	barely beats	C	A	destroys	C	0.929	0.397	0.175	+1.35	-0.33	-1.03
A	destroys	B	B	beats	C	A	barely beats	C	0.783	0.451	0.266	+1.32	-0.23	-1.09
A	beats	B	B	barely beats	C	A	barely beats	C	0.739	0.461	0.300	+1.32	-0.21	-1.10
A	destroys	B	B	beats	C	A	beats	C	0.847	0.447	0.206	+1.31	-0.20	-1.11
A	beats	B	B	barely beats	C	A	beats	C	0.794	0.458	0.248	+1.31	-0.19	-1.12
A	destroys	B	B	beats	C	A	destroys	C	0.929	0.442	0.129	+1.30	-0.18	-1.13
A	beats	B	B	barely beats	C	A	destroys	C	0.864	0.455	0.182	+1.30	-0.16	-1.14
A	destroys	B	B	destroys	C	A	destroys	C	0.929	0.500	0.071	+1.22	0	-1.22
A	beats	B	B	beats	C	A	destroys	C	0.867	0.500	0.133	+1.22	0	-1.22
A	destroys	B	B	destroys	C	A	beats	C	0.844	0.500	0.156	+1.22	0	-1.22
A	barely beats	B	B	barely beats	C	A	destroys	C	0.813	0.500	0.188	+1.22	0	-1.22
A	beats	B	B	beats	C	A	beats	C	0.794	0.500	0.206	+1.22	0	-1.22
A	destroys	B	B	destroys	C	A	barely beats	C	0.778	0.500	0.222	+1.22	0	-1.22
A	barely beats	B	B	barely beats	C	A	beats	C	0.750	0.500	0.250	+1.22	0	-1.22
A	beats	B	B	beats	C	A	barely beats	C	0.737	0.500	0.263	+1.22	0	-1.22
A	barely beats	B	B	barely beats	C	A	barely beats	C	0.700	0.500	0.300	+1.22	0	-1.22
A	barely beats	B	B	beats	C	A	destroys	C	0.818	0.545	0.136	+1.14	+0.16	-1.30
A	beats	B	B	destroys	C	A	destroys	C	0.871	0.558	0.071	+1.13	+0.18	-1.30
A	barely beats	B	B	beats	C	A	beats	C	0.752	0.542	0.206	+1.12	+0.19	-1.31
A	beats	B	B	destroys	C	A	beats	C	0.794	0.553	0.153	+1.11	+0.20	-1.31
A	barely beats	B	B	beats	C	A	barely beats	C	0.700	0.539	0.261	+1.10	+0.21	-1.32
A	beats	B	B	destroys	C	A	barely beats	C	0.734	0.549	0.217	+1.09	+0.23	-1.32
A	barely beats	B	B	destroys	C	A	destroys	C	0.825	0.603	0.071	+1.03	+0.33	-1.35
A	barely beats	B	B	destroys	C	A	beats	C	0.755	0.594	0.150	+1.00	+0.37	-1.37
A	barely beats	B	B	destroys	C	A	barely beats	C	0.700	0.588	0.213	+0.96	+0.42	-1.38

Now, consider the oldest dilemma in rankings: A beats B, B beats C, but C beat A. We can solve this analytically with what we have learned! In the cases where all 3 victories are by the same margin, we get $r_A = 0.5, r_B = 0.5, r_C = 0.5$ as expected. But if we have discrepancies in the margin of victory, we can analytically say who the best team is. It turns out the 3 equations are solved in greatest favor for team A when A destroys B, B barely beats C, and C barely beats A. In that case you have $r_A = 0.563$.

									rA	rB	rC	σA	σB	σC
A	destroys	B	B	beats	C	C	barely beats	A	0.559	0.465	0.476	+1.41	-0.83	-0.58
A	destroys	B	B	barely beats	C	C	barely beats	A	0.563	0.438	0.500	+1.22	-1.22	0
A	destroys	B	B	destroys	C	C	barely beats	A	0.556	0.500	0.444	+1.22	0	-1.22
A	destroys	B	B	beats	C	C	beats	A	0.533	0.467	0.500	+1.22	-1.22	0
A	destroys	B	B	destroys	C	C	beats	A	0.531	0.500	0.469	+1.22	0	-1.22
A	beats	B	B	barely beats	C	C	barely beats	A	0.528	0.472	0.500	+1.22	-1.22	0
A	beats	B	B	beats	C	C	barely beats	A	0.526	0.500	0.474	+1.22	0	-1.22
A	destroys	B	B	barely beats	C	C	beats	A	0.535	0.441	0.524	+0.83	-1.41	+0.58
A	beats	B	B	destroys	C	C	barely beats	A	0.524	0.535	0.441	+0.58	+0.83	-1.41
A	barely beats	B	B	destroys	C	C	barely beats	A	0.500	0.563	0.438	0	+1.22	-1.22
A	beats	B	B	destroys	C	C	beats	A	0.500	0.533	0.467	0	+1.22	-1.22
A	barely beats	B	B	beats	C	C	barely beats	A	0.500	0.528	0.472	0	+1.22	-1.22
A	destroys	B	B	destroys	C	C	destroys	A	0.500	0.500	0.500	0	0	0
A	barely beats	B	B	barely beats	C	C	barely beats	A	0.500	0.500	0.500	0	0	0
A	beats	B	B	beats	C	C	beats	A	0.500	0.500	0.500	0	0	0
A	beats	B	B	barely beats	C	C	beats	A	0.500	0.474	0.526	0	-1.22	+1.22
A	destroys	B	B	beats	C	C	destroys	A	0.500	0.469	0.531	0	-1.22	+1.22
A	destroys	B	B	barely beats	C	C	destroys	A	0.500	0.444	0.556	0	-1.22	+1.22
A	barely beats	B	B	destroys	C	C	beats	A	0.476	0.559	0.465	-0.58	+1.41	-0.83
A	beats	B	B	barely beats	C	C	destroys	A	0.465	0.476	0.559	-0.83	-0.58	+1.41
A	barely beats	B	B	beats	C	C	beats	A	0.474	0.526	0.500	-1.22	+1.22	0
A	barely beats	B	B	barely beats	C	C	beats	A	0.472	0.500	0.528	-1.22	0	+1.22
A	beats	B	B	destroys	C	C	destroys	A	0.469	0.531	0.500	-1.22	+1.22	0
A	beats	B	B	beats	C	C	destroys	A	0.467	0.500	0.533	-1.22	0	+1.22
A	barely beats	B	B	destroys	C	C	destroys	A	0.444	0.556	0.500	-1.22	+1.22	0
A	barely beats	B	B	barely beats	C	C	destroys	A	0.438	0.500	0.563	-1.22	0	+1.22
A	barely beats	B	B	beats	C	C	destroys	A	0.441	0.524	0.535	-1.41	+0.58	+0.83

The classic case of A>B>C>A dilemma is the 2008 Big 12 south season between Oklahoma, Texas, and Texas Tech. During that season Texas beat Oklahoma in the Cotton Bowl by 10 points. Later in the season Oklahoma destroyed Texas Tech by 44 points. However, a week before that drubbing Texas Tech barely beat Texas on a last second pass by 6 points. All three teams ended the regular season with identical 11-1 records, the only blemishes were to each other.

In the final AP poll before the conference championship week 8 more voters decided to elevate Texas as the best group of that trio, while the coaches' poll had Oklahoma on top. Highlighted in blue in table X is the Texas-Oklahoma-Texas Tech scenario, with Texas as team A, Oklahoma as team B, and Texas Tech as team C. Our equations are solved in this case in favor of team B, Oklahoma.

Other interesting things fall out due to the solution of the simultaneous equations. Consider a top team that has gone 12-0. This team has beaten both a top 10 team by 1 point, and a bottom 10 team, also by 1 point. As margin is increased from a 1-point victory to a M point victory for the game against the bottom 10 team, the rating of the 12-0 team increases approximately 2X how its rating would increase for the same M point victory against the top 10 team (returning the bottom 10 team victory to 1-point). What is happening is that the probability space from a bottom team to a top team should be very large and a 1-point victory there is much less likely to be explained by the left tail of one distribution matching up to the right tail of another distribution, than you can explain when two teams with similar distributions and one selects the middle, and the other selects the tail. In short, you are punished more for not destroying a bottom tier team than you are rewarded for destroying a top tier team. It may be slightly counterintuitive, but it makes perfect sense to the math.

4.0 Normalization of ratings

The addition of MoV adds an element that makes the range of ratings as much as (+1.5,-5) by the end of a complete FBS season. Some top teams are just better and amass a bunch of points on the field. Recalling that the average rating is $r = 0.5$, the ratings can be

- converted to Z-scores, or
- normalize to around (0-1) by use of $r_{final} = 0.5 + \ (r_{initial} - 0.5) \ * \ (1 - \alpha)$

At this time the playoffPredictor.com model does not normalize the ratings.

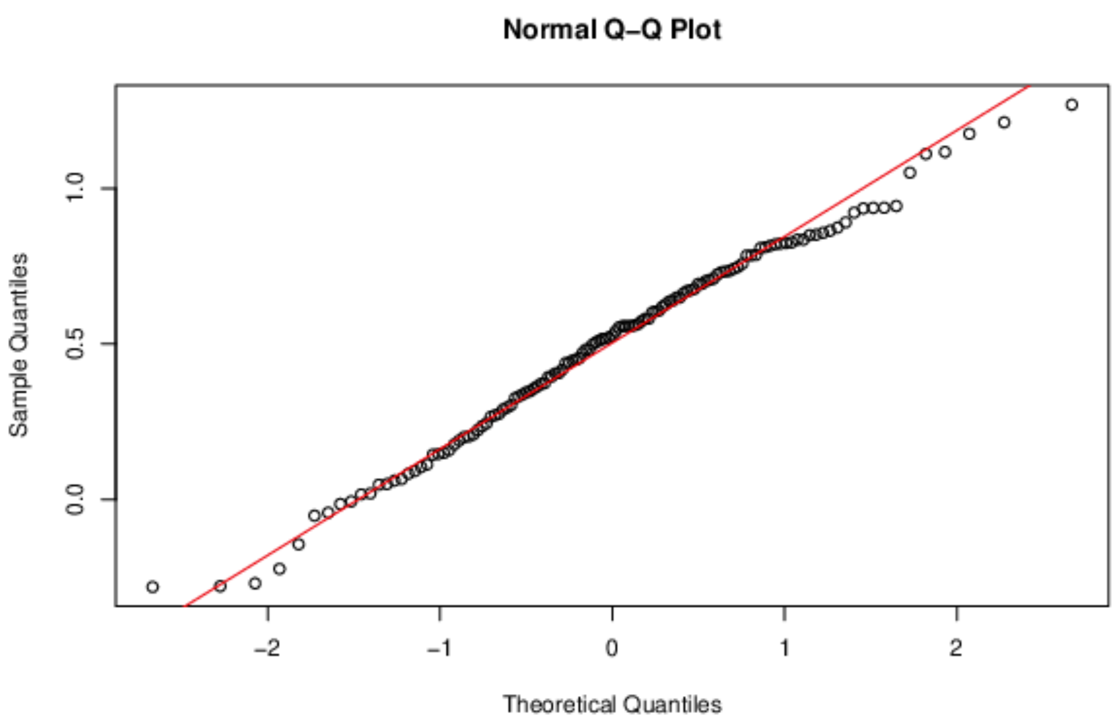
Note, this is purely a transformation done for aesthetic purposes only. Converting to Z-scores or normalizing to a range of (0,1) will still keep the same teams in the same order and the same relative spacing between each, just stretched or compressed.

Normal distribution of the ratings

Note that the ratings are a normal distribution. This is seen in a Q-Q plot of ratings:

We can test this with a Normal Q-Q plot. Below we show the week 14 $\vec{r}$ for all teams in 2021 with the actual 2021 $\vec{r}$ plotted against theoretical:

- Future predictions based on 0 movement of ratings (2022 week 10 Oregon over Colorado by 30)



Elo-ness of the model

This model is also Elo-like in the sense it is a zero sum model. No one team can increase in rating without a matching decrease coming from the rest of the pool. The sum of all teams ratings is always 0.5 * # teams.

5.0 moving the model to 5 teams

Consider a 5-team league, teams A-E. The league plays a total of 7 games with the following results:

E	1	C	0
E	0	D	1
E	0	A	1
B	0	C	1
B	1	A	0
C	1	D	0
C	0	A	1

team	A	B	C	D	E		record	initial rating w/o strength of schedule
A		L	W		W		2 - 1	0.6
B	W		L				1 - 1	0.5
C	L	W		W	L		2 - 2	0.5
D			L		W		1 - 1	0.5
E	L		W	L			1 - 2	0.4

Figure 2: 5 team league results. Read entries left to right (not top to bottom) - A beats C and E, A loses to B

In our world these are all 1-point wins by the score of 1-0[7] so expressly as to remove margin of victory from consideration (but not strength of schedule). Note that we have a seemingly better team "A" with a record of 2-1 and a worst team "E" with a record of 1-2. B, C, D all have .500 records, with B and D having played 2 games and C played 4 games – once against each member of the league. How would you go about ranking these teams?

After inputting strength-of-schedule with the original formula (no MoV) we come to the following ratings:

Final standings	final rating
A	0.587
B	0.522
C	0.500
D	0.478
E	0.413

Team B is clearly better than team D, because it's one win is against A, the best team in the league, where team D has only a win against team E, the worst team in the league. Now let's look at what happens if team A's wins over C and E are dominant wins, keeping a 1-point loss to team B.

[7] 1-0 is in fact not a normal nor possible football score. There is a concept of a 1-point safety, but it would have to follow a touchdown. Therefore, a final score of 6-1 is possible.

E	1	C	0
E	0	D	1
E	0	A	40
B	0	C	1
B	1	A	0
C	1	D	0
C	0	A	20

Final standings	final rating
A	0.692
B	0.542
C	0.474
D	0.453
E	0.339

Note that A gets significantly stronger, moving up +.105, at the expense of -.026 from C and -.074 from E. Where did that extra +.005 come from? It came from D, who dropped a whole -.025. The extra +0.2 went to B, who got rewarded for beating C and barely losing to A.

In fact, Margin of Victory can cause changes to the order. Imagine this scenario:

E	1	C	0
E	0	D	1
E	0	A	1
B	0	C	1
B	40	A	0
C	1	D	0
C	0	A	1

Final standings	final rating
B	0.578
A	0.543
C	0.500
D	0.476
E	0.404

While all other victories have been reset to 1 point, B’s win over A was dominant – 40 points (for this system we have been using m linear, 40 points equals $m = 1$, where 20 points results in $m = 0.5$). This forces B to be on top to satisfy the 5 equations with 5 unknowns. In this particular model B needs a 28-point victory $m = 0.7$ to have a higher final rating (at that point they both get tied at a .558 rating).

6.0 model applied to a full FBS season

Let's model the 2022 season and accuracy of this method to the human polls at 4 points in time: Week 5, Week 9, Week 13, and Week 14. For each week we will compare the computer method to the AP/committee ranking for that week and the committee final ranking for $\alpha = 0$ (which is the Colley method without the groups of FCS teams) and $\alpha = 0.5$ (which is the playofPredictor.com method)

6.1 computing η

Eta definition

Midseason

Final season

Convergence without MoV

Convergence with Mov

(Comparisons start on the next page)

Week 5:

Green cells are for user input	
α	0
yellow cells are calculated	
mean rating	0.500
std deviation	0.174
Aii adjustment	0
Aij adjustment	0
Dii adjustment (w)	0
~~not used~~	~~0~~
~~not used~~	~~3~~
~~not used~~	~~1~~

- Negative numbers mean team is underrated by committee
+ Positive numbers mean team is overrated by committee

		team	rating	Z*-score	Z* range 1-10
(5-0)	1	Ohio State	0.8414	+1.959	
(5-0)	2	Alabama	0.8407	+1.955	
(5-0)	3	Clemson	0.8276	+1.880	
(5-0)	4	Kansas	0.8075	+1.764	
(5-0)	5	Georgia	0.8018	+1.732	
(5-0)	6	Ole Miss	0.8006	+1.725	
(5-0)	7	Penn State	0.7956	+1.696	
(5-0)	8	USC	0.7884	+1.655	
(5-0)	9	UCLA	0.7826	+1.622	
(4-0)	10	Tennessee	0.7776	+1.593	0.366
(5-0)	11	Michigan	0.7763	+1.586	
(5-0)	12	Syracuse	0.7664	+1.529	
(4-1)	13	Mississippi State	0.7661	+1.527	
(4-1)	14	Wake Forest	0.7561	+1.470	
(4-0)	15	James Madison	0.7480	+1.423	
(4-0)	16	TCU	0.7277	+1.307	
(4-1)	17	Oregon	0.7222	+1.275	
(4-1)	18	LSU	0.7145	+1.231	
(5-0)	19	Coastal Carolina	0.7044	+1.173	
(4-0)	20	Oklahoma State	0.7040	+1.171	
(4-1)	21	Liberty	0.6945	+1.116	
(4-1)	22	Florida State	0.6926	+1.105	
(4-1)	23	BYU	0.6856	+1.065	
(4-1)	24	NC State	0.6839	+1.055	
(4-1)	25	Memphis	0.6835	+1.053	
(4-1)	26	Kansas State	0.6831	+1.051	
(4-1)	27	Maryland	0.6736	+0.996	
(4-1)	28	Washington	0.6537	+0.882	
(4-1)	29	Kentucky	0.6392	+0.799	
(4-1)	30	Tulane	0.6331	+0.764	
(4-1)	31	Duke	0.6310	+0.752	
(4-1)	32	Washington State	0.6294	+0.742	

Committee/Poll Week	5	Diff to Aga-elo mat	$\lvert \ln A(i) - \ln C(i) \rvert$
1	Alabama	1	0.693
2	Georgia	3	0.916
3	Ohio State	-2	1.099
4	Michigan	7	1.012
5	Clemson	-2	0.511
6	USC	2	0.288
7	Oklahoma State	13	1.050
8	Tennessee	2	0.223
9	Ole Miss	-3	0.405
10	Penn State	-3	0.357
11	Utah	25	1.186
12	Oregon	5	0.348
13	Kentucky	16	0.802
14	NC State	10	0.539
15	Wake Forest	-1	0.069
16	BYU	7	0.363
17	TCU	-1	0.061
18	UCLA	-9	0.693
19	Kansas	-15	1.558
20	Kansas State	6	0.262
21	Washington	7	0.288
22	Syracuse	-10	0.606
23	Mississippi State	-10	0.571
24	Cincinnati	10	0.348
25	LSU	-7	0.329
		calculated η →	1.791
	$\eta(\alpha)$	η improvement	
η (0)	1.791	-	

Committee/Poll Week	14	Diff to Aga-elo mat	$\lvert \ln A(i) - \ln C(i) \rvert$
1	Georgia	4	1.609
2	Michigan	9	1.705
3	TCU	13	1.674
4	Ohio State	-3	1.386
5	Alabama	-3	0.916
6	Tennessee	4	0.511
7	Clemson	-4	0.847
8	Utah	28	1.504
9	Kansas State	17	1.061
10	USC	-2	0.223
11	Penn State	-4	0.452
12	Washington	16	0.847
13	Florida State	9	0.526
14	Oregon State	43	1.404
15	Oregon	2	0.125
16	Tulane	14	0.629
17	LSU	1	0.057
18	UCLA	-9	0.693
19	South Carolin	52	1.318
20	Texas	19	0.668
21	Notre Dame	25	0.784
22	Mississippi St	-9	0.526
23	NC State	1	0.043
24	Troy	29	0.792
25	UTSA	43	1.001
		calculated η →	2.345
	$\eta(\alpha)$	η improvement	
η (0)	2.345	-	

Figure 3: 2022 week 5 - alpha=0 (no Margin of Victory Considered)

Green cells are for user input	
α	0.5
yellow cells are calculated	
mean rating	0.500
std deviation	0.217
Aii adjustment	0.5
Aij adjustment	0.5
Dii adjustment (w)	0
~~not used~~	~~0~~
~~not used~~	~~3~~
~~not used~~	~~1~~

- Negative numbers mean team is underrated by committee
+ Positive numbers mean team is overrated by committee

		team	rating	Z*-score	Z* range 1-10
(5-0)	1	Alabama	1.0089	+2.342	
(5-0)	2	Ohio State	1.0043	+2.321	
(5-0)	3	Georgia	0.9148	+1.909	
(5-0)	4	Clemson	0.8916	+1.802	
(5-0)	5	Michigan	0.8842	+1.768	
(5-0)	6	Ole Miss	0.8821	+1.758	
(5-0)	7	USC	0.8684	+1.696	
(5-0)	8	Penn State	0.8681	+1.694	
(5-0)	9	Kansas	0.8671	+1.690	
(4-0)	10	Tennessee	0.8656	+1.682	0.660
(4-1)	11	Mississippi State	0.8402	+1.566	
(5-0)	12	UCLA	0.8357	+1.545	
(4-0)	13	James Madison	0.8307	+1.522	
(5-0)	14	Syracuse	0.8296	+1.517	
(4-0)	15	TCU	0.8029	+1.394	
(4-1)	16	Wake Forest	0.7937	+1.352	
(4-1)	17	LSU	0.7871	+1.321	
(4-0)	18	Oklahoma State	0.7560	+1.178	
(4-1)	19	Oregon	0.7368	+1.090	
(4-1)	20	Kansas State	0.7281	+1.050	
(4-1)	21	Florida State	0.7251	+1.036	
(4-1)	22	Maryland	0.7243	+1.032	
(4-1)	23	NC State	0.7222	+1.023	
(5-0)	24	Coastal Carolina	0.7077	+0.956	
(4-1)	25	Liberty	0.7068	+0.952	
(4-1)	26	Washington	0.6950	+0.898	
(4-1)	27	Memphis	0.6941	+0.893	
(4-1)	28	BYU	0.6909	+0.878	
(4-1)	29	Utah	0.6815	+0.835	
(4-1)	30	Illinois	0.6756	+0.808	
(3-2)	31	Texas	0.6738	+0.800	
(4-1)	32	Kentucky	0.6715	+0.789	

Committee/Poll Week	5	Diff to Aga-elo mat	$\lvert \ln A(i) - \ln C(i) \rvert$
1	Alabama	0	0
2	Georgia	1	0.405
3	Ohio State	-1	0.405
4	Michigan	1	0.223
5	Clemson	-1	0.223
6	USC	1	0.154
7	Oklahoma State	11	0.944
8	Tennessee	2	0.223
9	Ole Miss	-3	0.405
10	Penn State	-2	0.223
11	Utah	18	0.969
12	Oregon	7	0.460
13	Kentucky	19	0.901
14	NC State	9	0.496
15	Wake Forest	1	0.065
16	BYU	12	0.560
17	TCU	-2	0.125
18	UCLA	-6	0.405
19	Kansas	-10	0.747
20	Kansas State	0	0
21	Washington	5	0.214
22	Syracuse	-8	0.452
23	Mississippi State	-12	0.738
24	Cincinnati	9	0.318
25	LSU	-8	0.386
		calculated η →	1.494
	$\eta(\alpha)$	η improvement	
η (0)	1.791	-	
η (.5)	1.494	37%	

Committee/Poll Week	14	Diff to Aga-elo mat	$\lvert \ln A(i) - \ln C(i) \rvert$
1	Georgia	2	1.099
2	Michigan	3	0.916
3	TCU	12	1.609
4	Ohio State	-2	0.693
5	Alabama	-4	1.609
6	Tennessee	4	0.511
7	Clemson	-3	0.560
8	Utah	21	1.288
9	Kansas State	11	0.799
10	USC	-3	0.357
11	Penn State	-3	0.318
12	Washington	14	0.773
13	Florida State	8	0.480
14	Oregon State	44	1.421
15	Oregon	4	0.236
16	Tulane	18	0.754
17	LSU	0	0
18	UCLA	-6	0.405
19	South Carolin	55	1.360
20	Texas	11	0.438
21	Notre Dame	21	0.693
22	Mississippi St	-11	0.693
23	NC State	0	0
24	Troy	31	0.829
25	UTSA	40	0.956
		calculated η →	2.121
	$\eta(\alpha)$	η improvement	
η (0)	2.345	-	
η (.5)	2.121	17%	

Figure 4: 2022 week 5 - alpha =0.5

Note that we get an improvement in η of 37% to the week 5 poll, and an improvement of 17% to the final poll (which was of course unknowable at the time of week 5). All of the top 6 teams in the computer ranking are the same top 6 teams in the AP poll, with no one team more than 1 position off its poll ranking. Without MoV Michigan alone is 7 spots off.

Week 9:

Green cells are for user input	
α	0
yellow cells are calculated	
mean rating	0.500
std deviation	0.188
Aii adjustment	0
Aij adjustment	0
Dii adjustment (w)	0
~~not used~~	~~0~~
~~not used~~	~~3~~
~~not used~~	~~1~~

- Negative numbers mean team is underrated by committee
+ Positive numbers mean team is overrated by committee

		team	rating	Z*-score	Z* range 1-10
(8-0)	1	Ohio State	0.8972	+2.117	
(8-0)	2	TCU	0.8943	+2.101	
(8-0)	3	Clemson	0.8867	+2.061	
(8-0)	4	Tennessee	0.8812	+2.031	
(8-0)	5	Georgia	0.8770	+2.009	
(8-0)	6	Michigan	0.8540	+1.886	
(7-1)	7	Alabama	0.8299	+1.758	
(7-1)	8	USC	0.8130	+1.668	
(7-1)	9	Oregon	0.8028	+1.613	
(7-1)	10	UCLA	0.8004	+1.601	0.516
(8-1)	11	Ole Miss	0.7911	+1.551	
(6-2)	12	Kansas State	0.7523	+1.344	
(7-1)	13	Tulane	0.7427	+1.293	
(6-2)	14	Syracuse	0.7358	+1.256	
(6-2)	15	LSU	0.7327	+1.240	
(7-1)	16	North Carolina	0.7318	+1.235	
(6-2)	17	Utah	0.7271	+1.210	
(6-2)	18	Penn State	0.7260	+1.205	
(7-1)	19	Illinois	0.7244	+1.196	
(6-2)	20	NC State	0.7163	+1.152	
(6-2)	21	Oklahoma State	0.7139	+1.140	
(7-1)	22	Coastal Carolina	0.7024	+1.078	
(6-2)	23	Wake Forest	0.6988	+1.059	
(6-2)	24	Oregon State	0.6941	+1.034	
(7-1)	25	Liberty	0.6896	+1.010	
(6-2)	26	Troy	0.6820	+0.970	
(6-2)	27	Maryland	0.6696	+0.904	
(5-3)	28	Notre Dame	0.6606	+0.856	
(5-3)	29	Mississippi State	0.6545	+0.823	
(5-3)	30	Texas	0.6482	+0.790	
(5-3)	31	Florida State	0.6452	+0.774	
(5-3)	32	Oklahoma	0.6404	+0.748	

Committee/Poll Week 9		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
1	Tennessee	3	1.386
2	Ohio State	-1	0.693
3	Georgia	2	0.511
4	Clemson	-1	0.288
5	Michigan	1	0.182
6	Alabama	1	0.154
7	TCU	-5	1.253
8	Oregon	1	0.118
9	USC	-1	0.118
10	LSU	5	0.405
11	Ole Miss	0	0
12	UCLA	-2	0.182
13	Kansas State	-1	0.080
14	Utah	3	0.194
15	Penn State	3	0.182
16	Illinois	3	0.172
17	North Carolina	-1	0.061
18	Oklahoma State	3	0.154
19	Tulane	-6	0.379
20	Syracuse	-6	0.357
21	Wake Forest	2	0.091
22	NC State	-2	0.095
23	Oregon State	1	0.043
24	Texas	6	0.223
25	UCF	9	0.307
		calculated η →	1.357

Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
1	Georgia	4	1.609
2	Michigan	4	1.099
3	TCU	-1	0.405
4	Ohio State	-3	1.386
5	Alabama	2	0.336
6	Tennessee	-2	0.405
7	Clemson	-4	0.847
8	Utah	9	0.754
9	Kansas State	3	0.288
10	USC	-2	0.223
11	Penn State	7	0.492
12	Washington	23	1.070
13	Florida State	18	0.869
14	Oregon State	10	0.539
15	Oregon	-6	0.511
16	Tulane	-3	0.208
17	LSU	-2	0.125
18	UCLA	-8	0.588
19	South Carolin	27	0.884
20	Texas	10	0.405
21	Notre Dame	7	0.288
22	Mississippi St	7	0.276
23	NC State	-3	0.140
24	Troy	2	0.080
25	UTSA	8	0.278
		calculated η →	1.758

	η (α)	η improvement			η (α)	η improvement
η (0)	1.357	-		η (0)	1.758	-

Figure 5: 2022 week 9 - no MoV

Green cells are for user input	
α	0.5
yellow cells are calculated	
mean rating	0.500
std deviation	0.234
Aii adjustment	0.5
Aij adjustment	0.5
Dii adjustment (w)	0
~~not used~~	~~0~~
~~not used~~	~~3~~
~~not used~~	~~1~~

- Negative numbers mean team is underrated by committee
+ Positive numbers mean team is overrated by committee

		team	rating	Z*-score	Z* range 1-10
(8-0)	1	Ohio State	1.0915	+2.532	
(8-0)	2	Georgia	1.0470	+2.342	
(8-0)	3	Tennessee	1.0335	+2.284	
(8-0)	4	Michigan	0.9930	+2.110	
(7-1)	5	Alabama	0.9823	+2.064	
(8-0)	6	TCU	0.9803	+2.056	
(8-0)	7	Clemson	0.9612	+1.974	
(7-1)	8	USC	0.8937	+1.685	
(7-1)	9	UCLA	0.8580	+1.532	
(7-1)	10	Oregon	0.8537	+1.514	1.018
(8-1)	11	Ole Miss	0.8413	+1.461	
(6-2)	12	Kansas State	0.8323	+1.422	
(6-2)	13	LSU	0.7980	+1.276	
(6-2)	14	Utah	0.7958	+1.266	
(7-1)	15	Tulane	0.7948	+1.262	
(6-2)	16	Syracuse	0.7942	+1.259	
(6-2)	17	Penn State	0.7920	+1.250	
(7-1)	18	Illinois	0.7834	+1.213	
(7-1)	19	North Carolina	0.7575	+1.102	
(6-2)	20	Oklahoma State	0.7447	+1.047	
(6-2)	21	NC State	0.7442	+1.045	
(5-3)	22	Texas	0.7372	+1.015	
(6-2)	23	Wake Forest	0.7296	+0.983	
(6-2)	24	Oregon State	0.7181	+0.934	
(6-2)	25	Maryland	0.7115	+0.905	
(5-3)	26	Notre Dame	0.7065	+0.884	
(7-1)	27	Liberty	0.7040	+0.873	
(5-3)	28	Mississippi State	0.7015	+0.863	
(7-1)	29	Coastal Carolina	0.6947	+0.833	
(6-2)	30	Troy	0.6912	+0.818	
(6-2)	31	UCF	0.6847	+0.791	
(5-3)	32	Florida State	0.6805	+0.773	

Committee/Poll Week 9		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
1	Tennessee	2	1.099
2	Ohio State	-1	0.693
3	Georgia	-1	0.405
4	Clemson	3	0.560
5	Michigan	-1	0.223
6	Alabama	-1	0.182
7	TCU	-1	0.154
8	Oregon	2	0.223
9	USC	-1	0.118
10	LSU	3	0.262
11	Ole Miss	0	0
12	UCLA	-3	0.288
13	Kansas State	-1	0.080
14	Utah	0	0
15	Penn State	2	0.125
16	Illinois	2	0.118
17	North Carolina	2	0.111
18	Oklahoma State	2	0.105
19	Tulane	-4	0.236
20	Syracuse	-4	0.223
21	Wake Forest	2	0.091
22	NC State	-1	0.047
23	Oregon State	1	0.043
24	Texas	-2	0.087
25	UCF	6	0.215
		calculated η →	1.256

Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
1	Georgia	1	0.693
2	Michigan	2	0.693
3	TCU	3	0.693
4	Ohio State	-3	1.386
5	Alabama	0	0
6	Tennessee	-3	0.693
7	Clemson	0	0
8	Utah	6	0.560
9	Kansas State	3	0.288
10	USC	-2	0.223
11	Penn State	6	0.435
12	Washington	21	1.012
13	Florida State	19	0.901
14	Oregon State	10	0.539
15	Oregon	-5	0.405
16	Tulane	-1	0.065
17	LSU	-4	0.268
18	UCLA	-9	0.693
19	South Carolin	31	0.968
20	Texas	2	0.095
21	Notre Dame	5	0.214
22	Mississippi St	6	0.241
23	NC State	-2	0.091
24	Troy	6	0.223
25	UTSA	12	0.392
		calculated η →	1.601

	η (α)	η improvement			η (α)	η improvement
η (0)	1.357	-		η (0)	1.758	-
η (.5)	1.256	28%		η (.5)	1.601	21%

Figure 6: 2022 week 9 - alpha = 0.5

For week 9 we get an improvement in η of 28% to the week 9 poll, and an improvement of 21% to the final poll.

Now we look at week 13, which is the last regular-season week:

Green cells are for user input		Aii adjustment	0
α	0	Aij adjustment	0
mean rating	0.500	Dii adjustment (w)	0
std deviation	0.195	~~not used~~	~~0~~
yellow cells are calculated		~~not used~~	~~3~~
		~~not used~~	~~1~~

- Negative numbers mean team is underrated by committee
+ Positive numbers mean team is overrated by committee

		team	rating	Z*-score	Z* range 1-10
(12-0)	1	Georgia	0.9762	+2.440	
(12-0)	2	TCU	0.9456	+2.284	
(12-0)	3	Michigan	0.9196	+2.150	
(11-1)	4	USC	0.8789	+1.942	
(11-1)	5	Ohio State	0.8692	+1.892	
(10-2)	6	Tennessee	0.8319	+1.701	
(10-2)	7	Alabama	0.8267	+1.674	
(10-2)	8	Penn State	0.8149	+1.614	
(10-2)	9	Clemson	0.7941	+1.507	
(9-3)	10	Kansas State	0.7740	+1.404	1.036
(9-3)	11	Oregon	0.7635	+1.350	
(10-2)	12	Washington	0.7610	+1.337	
(9-3)	13	LSU	0.7532	+1.298	
(8-4)	14	Texas	0.7428	+1.244	
(9-3)	15	Oregon State	0.7415	+1.238	
(9-3)	16	Florida State	0.7394	+1.227	
(9-3)	17	Utah	0.7351	+1.205	
(10-2)	18	Tulane	0.7307	+1.182	
(9-3)	19	UCLA	0.7280	+1.169	
(10-2)	20	Troy	0.7234	+1.145	
(8-4)	21	Mississippi State	0.7027	+1.039	
(8-4)	22	Notre Dame	0.6969	+1.009	
(10-2)	23	UTSA	0.6968	+1.009	
(8-4)	24	Ole Miss	0.6831	+0.938	
(10-2)	25	South Alabama	0.6792	+0.918	
(9-2)	26	Coastal Carolina	0.6782	+0.913	
(9-3)	27	North Carolina	0.6772	+0.908	
(8-4)	28	South Carolina	0.6733	+0.888	
(8-4)	29	NC State	0.6630	+0.835	
(9-3)	30	UCF	0.6529	+0.784	
(9-3)	31	Cincinnati	0.6404	+0.719	
(7-5)	32	Texas Tech	0.6361	+0.698	

Committee/Poll Week	13	Diff to Aga-elo mat	|ln A(i)−ln C(i)|
1	Georgia	0	0
2	Michigan	1	0.405
3	TCU	-1	0.405
4	USC	0	0
5	Ohio State	0	0
6	Alabama	1	0.154
7	Tennessee	-1	0.154
8	Penn State	0	0
9	Clemson	0	0
10	Kansas State	0	0
11	Utah	6	0.435
12	Washington	0	0
13	Florida State	3	0.208
14	LSU	-1	0.074
15	Oregon State	0	0
16	Oregon	-5	0.375
17	UCLA	2	0.111
18	Tulane	0	0
19	South Carolina	9	0.388
20	Texas	-6	0.357
21	Notre Dame	1	0.047
22	UCF	8	0.310
23	North Carolina	4	0.160
24	Mississippi State	-3	0.134
25	NC State	4	0.148
		calculated η →	1.167
	η (α)	η improvement	
η (0)	1.167	-	

Committee/Poll Week	14	Diff to Aga-elo mat	|ln A(i)−ln C(i)|
1	Georgia	0	0
2	Michigan	1	0.405
3	TCU	-1	0.405
4	Ohio State	1	0.223
5	Alabama	2	0.336
6	Tennessee	0	0
7	Clemson	2	0.251
8	Utah	9	0.754
9	Kansas State	1	0.105
10	USC	-6	0.916
11	Penn State	-3	0.318
12	Washington	0	0
13	Florida State	3	0.208
14	Oregon State	1	0.069
15	Oregon	-4	0.310
16	Tulane	2	0.118
17	LSU	-4	0.268
18	UCLA	1	0.054
19	South Carolin	9	0.388
20	Texas	-6	0.357
21	Notre Dame	1	0.047
22	Mississippi St	-1	0.047
23	NC State	6	0.232
24	Troy	-4	0.182
25	UTSA	-2	0.083
		calculated η →	1.275
	η (α)	η improvement	
η (0)	1.275	-	

Figure 7: 2022 week 13 - no MoV

Green cells are for user input		Aii adjustment	0.5
α	0.5	Aij adjustment	0.5
mean rating	0.500	Dii adjustment (w)	0
std deviation	0.244	~~not used~~	~~0~~
yellow cells are calculated		~~not used~~	~~3~~
		~~not used~~	~~1~~

- Negative numbers mean team is underrated by committee
+ Positive numbers mean team is overrated by committee

		team	rating	Z*-score	Z* range 1-10
(12-0)	1	Georgia	1.1751	+2.768	
(12-0)	2	Michigan	1.0894	+2.416	
(12-0)	3	TCU	1.0541	+2.272	
(11-1)	4	Ohio State	1.0306	+2.175	
(10-2)	5	Tennessee	0.9779	+1.959	
(10-2)	6	Alabama	0.9681	+1.919	
(11-1)	7	USC	0.9674	+1.916	
(10-2)	8	Penn State	0.9242	+1.739	
(9-3)	9	Kansas State	0.8693	+1.514	
(10-2)	10	Clemson	0.8566	+1.462	1.306
(8-4)	11	Texas	0.8538	+1.451	
(9-3)	12	LSU	0.8232	+1.325	
(9-3)	13	Utah	0.8225	+1.322	
(9-3)	14	Florida State	0.8219	+1.320	
(9-3)	15	Oregon	0.8173	+1.301	
(10-2)	16	Washington	0.8124	+1.281	
(10-2)	17	Tulane	0.7883	+1.182	
(9-3)	18	Oregon State	0.7882	+1.181	
(9-3)	19	UCLA	0.7704	+1.109	
(8-4)	20	Notre Dame	0.7504	+1.026	
(8-4)	21	Mississippi State	0.7486	+1.019	
(10-2)	22	Troy	0.7351	+0.964	
(8-4)	23	Ole Miss	0.7274	+0.932	
(10-2)	24	UTSA	0.7124	+0.871	
(10-2)	25	South Alabama	0.6963	+0.805	
(9-3)	26	UCF	0.6951	+0.800	
(8-4)	27	South Carolina	0.6920	+0.787	
(9-3)	28	North Carolina	0.6903	+0.780	
(8-4)	29	NC State	0.6820	+0.746	
(7-5)	30	Texas Tech	0.6654	+0.678	
(9-3)	31	Cincinnati	0.6650	+0.677	
(7-5)	32	Louisville	0.6636	+0.671	

Committee/Poll Week	13	Diff to Aga-elo mat	|ln A(i)−ln C(i)|
1	Georgia	0	0
2	Michigan	0	0
3	TCU	0	0
4	USC	3	0.560
5	Ohio State	-1	0.223
6	Alabama	0	0
7	Tennessee	-2	0.336
8	Penn State	0	0
9	Clemson	1	0.105
10	Kansas State	-1	0.105
11	Utah	2	0.167
12	Washington	4	0.288
13	Florida State	1	0.074
14	LSU	-2	0.154
15	Oregon State	3	0.182
16	Oregon	-1	0.065
17	UCLA	2	0.111
18	Tulane	-1	0.057
19	South Carolina	8	0.351
20	Texas	-9	0.598
21	Notre Dame	-1	0.049
22	UCF	4	0.167
23	North Carolina	5	0.197
24	Mississippi State	-3	0.134
25	NC State	4	0.148
		calculated η →	1.177
	η (α)	η improvement	
η (0)	1.167	-	
η (.5)	1.177	-6%	

Committee/Poll Week	14	Diff to Aga-elo mat	|ln A(i)−ln C(i)|
1	Georgia	0	0
2	Michigan	0	0
3	TCU	0	0
4	Ohio State	0	0
5	Alabama	1	0.182
6	Tennessee	-1	0.182
7	Clemson	3	0.357
8	Utah	5	0.486
9	Kansas State	0	0
10	USC	-3	0.357
11	Penn State	-3	0.318
12	Washington	4	0.288
13	Florida State	1	0.074
14	Oregon State	4	0.251
15	Oregon	0	0
16	Tulane	1	0.061
17	LSU	-5	0.348
18	UCLA	1	0.054
19	South Carolin	8	0.351
20	Texas	-9	0.598
21	Notre Dame	-1	0.049
22	Mississippi St	-1	0.047
23	NC State	6	0.232
24	Troy	-2	0.087
25	UTSA	-1	0.041
		calculated η →	1.191
	η (α)	η improvement	
η (0)	1.275	-	
η (.5)	1.191	31%	

Figure 8: 2022 week 13 - alpha = 0.5

For week 13 we get a retrogression in η of -6% to the week 13 poll (adding in MoV made things *worse*), but still an improvement of 21% to the final poll.

Finally, let’s consider week 14, which is the conclusion of championship week, where several rematches are played:

Green cells are for user input	
α	0
yellow cells are calculated	
mean rating	0.500
std deviation	0.195
Aii adjustment	0
Aij adjustment	0
Dii adjustment (w)	0
~~not used~~	0
~~not used~~	3
~~not used~~	1
- Negative nu	
+ Positive nur	
Committee/Poll Week	14

		team	rating	Z*-score	Z* range 1-10			Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(13-0)	1	Georgia	0.9947	+2.530		1	Georgia	0	0
(13-0)	2	Michigan	0.9321	+2.210		2	Michigan	0	0
(12-1)	3	TCU	0.9054	+2.074		3	TCU	0	0
(11-1)	4	Ohio State	0.8730	+1.908		4	Ohio State	0	0
(11-2)	5	USC	0.8423	+1.751		5	Alabama	2	0.336
(10-2)	6	Tennessee	0.8312	+1.694		6	Tennessee	0	0
(10-2)	7	Alabama	0.8253	+1.664		7	Clemson	1	0.134
(11-2)	8	Clemson	0.8178	+1.626		8	Utah	3	0.318
(10-3)	9	Kansas State	0.8152	+1.613		9	Kansas State	0	0
(10-2)	10	Penn State	0.8124	+1.598	0.932	10	USC	-5	0.693
(10-3)	11	Utah	0.7737	+1.400		11	Penn State	-1	0.095
(9-3)	12	Oregon	0.7667	+1.364		12	Washington	1	0.080
(10-2)	13	Washington	0.7606	+1.333		13	Florida State	5	0.325
(11-2)	14	Tulane	0.7584	+1.322		14	Oregon State	3	0.194
(11-2)	15	Troy	0.7523	+1.291		15	Oregon	-3	0.223
(8-4)	16	Texas	0.7435	+1.246		16	Tulane	-2	0.134
(9-3)	17	Oregon State	0.7411	+1.233		17	LSU	2	0.111
(9-3)	18	Florida State	0.7400	+1.228		18	UCLA	2	0.105
(9-4)	19	LSU	0.7358	+1.206		19	South Carolina	7	0.314
(9-3)	20	UCLA	0.7280	+1.166		20	Texas	-4	0.223
(11-2)	21	UTSA	0.7126	+1.088		21	Notre Dame	2	0.091
(8-4)	22	Mississippi State	0.7025	+1.036		22	Mississippi State	0	0
(8-4)	23	Notre Dame	0.6938	+0.991		23	NC State	4	0.160
(8-4)	24	Ole Miss	0.6838	+0.940		24	Troy	-9	0.470
(10-2)	25	South Alabama	0.6816	+0.929		25	UTSA	-4	0.174
(8-4)	26	South Carolina	0.6766	+0.904				calculated η →	1.182
(8-4)	27	NC State	0.6628	+0.833					
(9-4)	28	North Carolina	0.6513	+0.774					
(9-3)	29	Coastal Carolina	0.6478	+0.756					
(9-3)	30	Cincinnati	0.6394	+0.713			η (α)	η improvement	
(7-5)	31	Texas Tech	0.6359	+0.695		η (0)	1.182	-	
(9-4)	32	UCF	0.6268	+0.649					
(7-5)	33	Kentucky	[illegible]	[illegible]					

Figure 9: 2022 week 14 - no MoV

Green cells are for user input	
α	0.5
yellow cells are calculated	
mean rating	0.500
std deviation	0.245
Aii adjustment	0.5
Aij adjustment	0.5
Dii adjustment (w)	0
~~not used~~	0
~~not used~~	3
~~not used~~	1
- Negative n	
+ Positive nu	
Committee/Poll Week	14

		team	rating	Z*-score	Z* range 1-10			Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(13-0)	1	Georgia	1.2006	+2.863		1	Georgia	0	0
(13-0)	2	Michigan	1.1050	+2.472		2	Michigan	0	0
(11-1)	3	Ohio State	1.0363	+2.192		3	TCU	1	0.288
(12-1)	4	TCU	1.0093	+2.081		4	Ohio State	-1	0.288
(10-2)	5	Tennessee	0.9771	+1.950		5	Alabama	1	0.182
(10-2)	6	Alabama	0.9662	+1.905		6	Tennessee	-1	0.182
(10-2)	7	Penn State	0.9208	+1.719		7	Clemson	3	0.357
(11-2)	8	USC	0.9171	+1.704		8	Utah	3	0.318
(10-3)	9	Kansas State	0.9150	+1.696		9	Kansas State	0	0
(11-2)	10	Clemson	0.8925	+1.604	1.259	10	USC	-2	0.223
(10-3)	11	Utah	0.8767	+1.539		11	Penn State	-4	0.452
(8-4)	12	Texas	0.8551	+1.451		12	Washington	4	0.288
(11-2)	13	Tulane	0.8240	+1.324		13	Florida State	1	0.074
(9-3)	14	Florida State	0.8231	+1.320		14	Oregon State	4	0.251
(9-3)	15	Oregon	0.8217	+1.315		15	Oregon	0	0
(10-2)	16	Washington	0.8120	+1.275		16	Tulane	-3	0.208
(9-4)	17	LSU	0.8012	+1.231		17	LSU	0	0
(9-3)	18	Oregon State	0.7871	+1.173		18	UCLA	2	0.105
(11-2)	19	Troy	0.7707	+1.106		19	South Carolina	7	0.314
(9-3)	20	UCLA	0.7707	+1.106		20	Texas	-8	0.511
(8-4)	21	Mississippi State	0.7483	+1.015		21	Notre Dame	1	0.047
(8-4)	22	Notre Dame	0.7463	+1.007		22	Mississippi State	-1	0.047
(11-2)	23	UTSA	0.7336	+0.955		23	NC State	4	0.160
(8-4)	24	Ole Miss	0.7286	+0.934		24	Troy	-5	0.234
(10-2)	25	South Alabama	0.6993	+0.814		25	UTSA	-2	0.083
(8-4)	26	South Carolina	0.6964	+0.802				calculated η →	1.203
(8-4)	27	NC State	0.6816	+0.742					
(7-5)	28	Texas Tech	0.6650	+0.674					
(9-3)	29	Cincinnati	0.6637	+0.669					
(7-5)	30	Louisville	0.6627	+0.665			η (α)	η improvement	
(9-4)	31	UCF	0.6608	+0.657		η (0)	1.182	-	
(9-4)	32	North Carolina	0.6524	+0.623		η (.5)	1.203	-11%	

Figure 10: 2022 week 14 - alpha =0.5

In the final poll itself we get a retrogression in η of -11%.

These results are typical -- η improvement of 20%-50% can be seen in early and midseason weeks, and η improvement goes anywhere from -10% to +10% by the end of the season.

Are these results good? Is the desire to have $\eta = 1$, which implies computer poll and AP/committee poll alignment? To the latter question, the answer is probably not. If computer ratings, even these simple ratings[8], discover hidden truths (like Texas A&M is not as good as the public thinks, or TCU is much better than people realize) early, than the computer formula will have provided a valuable service, and if this realization can be done with simple inputs of only final scores, it lends credence to composite ratings that incorporate more information than just scores.

We can also test other scenarios. If Oregon wins 9 games but loses one game by 46 points to Georgia, how bad is that one loss compared to if they lost a close game to Georgia by 3 points?

2022 ratings through week 14 (win perspective)

		team	rating	Z*-score
(13-0)	1	Georgia	1.2006	+2.863
(13-0)	2	Michigan	1.1050	+2.472
(11-1)	3	Ohio State	1.0363	+2.192
(12-1)	4	TCU	1.0093	+2.081
(10-2)	5	Tennessee	0.9771	+1.950
(10-2)	6	Alabama	0.9662	+1.905
(10-2)	7	Penn State	0.9208	+1.719
(11-2)	8	USC	0.9171	+1.704
(10-3)	9	Kansas State	0.9150	+1.696
(11-2)	10	Clemson	0.8925	+1.604
(10-3)	11	Utah	0.8767	+1.539
(8-4)	12	Texas	0.8551	+1.451
(11-2)	13	Tulane	0.8240	+1.324
(9-3)	14	Florida State	0.8231	+1.320
(9-3)	15	Oregon	0.8217	+1.315

Oregon is #15 which is the same as their final committee rank

Change week 1 to a final score of Georgia 6 - Oregon 3:

[8] I acknowledge the math may not seem simple, but it really is. Again, only 1 input is used, and we get good correlation with human polls and betting markets.

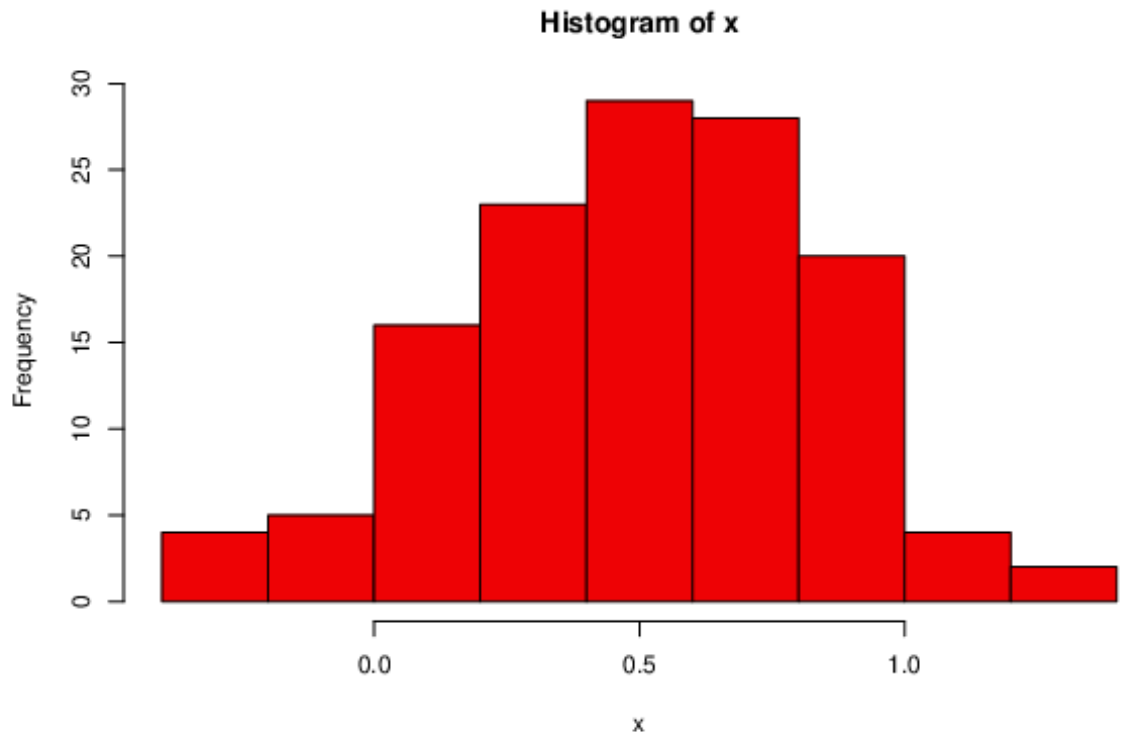
		team	rating	Z*-score
(13-0)	1	Georgia	1.1680	+2.733
(13-0)	2	Michigan	1.1051	+2.476
(11-1)	3	Ohio State	1.0365	+2.195
(12-1)	4	TCU	1.0095	+2.085
(10-2)	5	Tennessee	0.9724	+1.933
(10-2)	6	Alabama	0.9642	+1.900
(11-2)	7	USC	0.9207	+1.721
(10-2)	8	Penn State	0.9205	+1.721
(10-3)	9	Kansas State	0.9147	+1.697
(11-2)	10	Clemson	0.8917	+1.603
(10-3)	11	Utah	0.8821	+1.564
(9-3)	12	Oregon	0.8564	+1.458
(8-4)	13	Texas	0.8549	+1.452
(11-2)	14	Tulane	0.8238	+1.325
(9-3)	15	Florida State	0.8220	+1.317

Oregon moves from #15 to #12. Georgia's ranking stays #1, but their rating moves from 1.2006 to 1.1680.

From the loss perspective Oregon stayed #12, and Georgia went from 1.0420 to 1.0402

Washington (who had a bad loss to Arizona State) is the other example. Win perspective does not punish them as much as loss-perspective does.

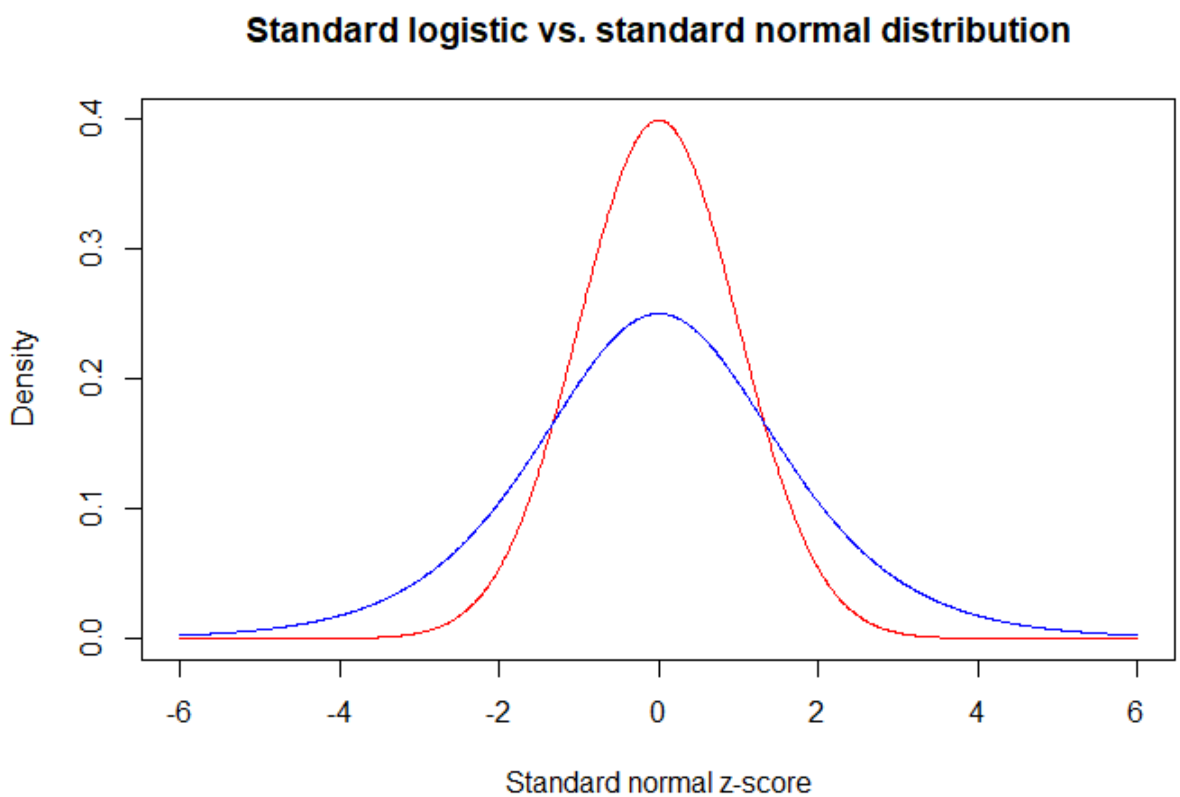
6.2 histogram of MoV for a season



We find in practice a good fit with exceptions for the $Z \geq \pm 2$ teams. Those teams are different in that they are the ones that play conference championship games and thus have different n_{tot} values than the other teams.

Let's define $\acute{r}_i$ as a random variable with mean r_i and standard deviation of 0.325. The league normal distribution is caused by *each team having a distribution with mean r_i*. Is $\acute{r}_i$ normal, and is its standard deviation also 0.325? Unfortunately, there is no way to know that theoretically. The Central Limit Theorem tells us that $\acute{r}_i$ as a random variable with *any* distribution and *any* standard deviation will cause the entire league $\vec{r}$ to be normally distributed. One thing we do have going for us – the Central Limit Theorem will only work if the standard deviations for all the $\acute{r}_i$'s are the same. This is different to what Elo thought for chess players. He posited that different players could have different standard deviations of their rating, but he made the simplifying assumption that every chess player's rating was a <u>normal</u> random variable with some mean and the same standard distribution. We cannot make any assumption based on theory of the type of distribution (normal, flat, logistic, etc) or the standard deviation of that distribution for each $\acute{r}_i$.

In chess it has been demonstrated that player's individual performance is better fitted as a logistic distributed random variable than as a normal distributed random variable. It is found that you need fatter tails because weaker players defeat higher ranked players more than the normal distribution would predict. Since, as stated in the preceding paragraph, we cannot know any theoretical basis for the distribution and variance of any given $\acute{r}_i$, let's do like chess and assume that all teams' true ratings are a random variable with mean r_i , standard deviation 0.325, and logistic distribution.



Pairwise probabilities from the logistic distribution can be mapped directly by

In this case we can then find probability for two teams matched together. The true win percentage for two teams $\acute{r}_A$ and $\acute{r}_B$ is based on $\sigma' = \sigma\sqrt{2}$ and given by $P(A) = \left[\frac{\acute{r}_B - \acute{r}_A}{0.325\sqrt{2}}\right] \sim N(0,1)$ This is a bad way to say the norm dist of rating diff divided by 0.325 * sqrt (2)

(whole league std dev or sqrt num games or sqrt num teams ?). then we can compute the probability as …

Suppose there is a game with three participants: Player A, Player B, and Player C. One player will finish in first place, another in second place, and another in third place (no ties allowed). I know the probability of all pairwise outcomes. For example, let's assume that Player A beats Player B 75% of the time, Player A beats Player C 90% of the time, and Player B beats Player C 75% of the time.

Given those pairwise probabilities, how can I calculate the probability of each possible outcome? With three players there are six possible outcomes: (A, B, C), (A, C, B), (B, A, C), (B, C, A), (C, A, B), or (C, B, A). I need to generalize to NN players for any $N \geq 3$N≥3. I'm ideally looking for an analytical answer, not a simulation.

If it helps, we can assume that the pairwise probabilities are derived from Elo ratings using a logistic distribution (description on Wikipedia if you're unfamiliar with Elo). This ensures that the pairwise probabilities are consistent (e.g., if A is likely to beat B and B is likely to beat C, then A is even more likely to beat C).

Here's what I ended up doing. My original question may have been a little ambiguous, but this solution gave me what I needed.

Rather than using the probabilities that Player A would beat Player B and so on, I assumed each player would have a "score" S_iSi sampled from some distribution D_iDi.
So $S_A \sim D_A$SA∼DA, $S_B \sim D_B$SB∼DB, and $S_C \sim D_C$SC∼DC, all independent, and then we sort the scores in descending order to see who gets first place, second place, and so on. Each of D_iDi is the same type of distribution but with different parameters. The challenge is to select a distribution and parameters that are consistent with the pairwise probabilities. That is, if we take the example pairwise probabilities from my original question, we want to select

distribution parameters such that $P(S_A > S_B) = 0.75$, $P(S_A > S_C) = 0.9$, and $P(S_B > S_C) = 0.75$.

In my case, the **Gumbel distribution** was the answer. That's because the difference of two Gumbel random variables follows a logistic distribution (see [Wikipedia]), and my pairwise probabilities are derived from the logistic function according to the standard Elo rating formula. (Technically, I used a modified version of the Gumbel distribution where I replaced the natural logs with log base 10).

To demonstrate, suppose I have three players with Elo ratings R_A, R_B, and R_C. Then according to Elo the probability of Player A beating Player B is

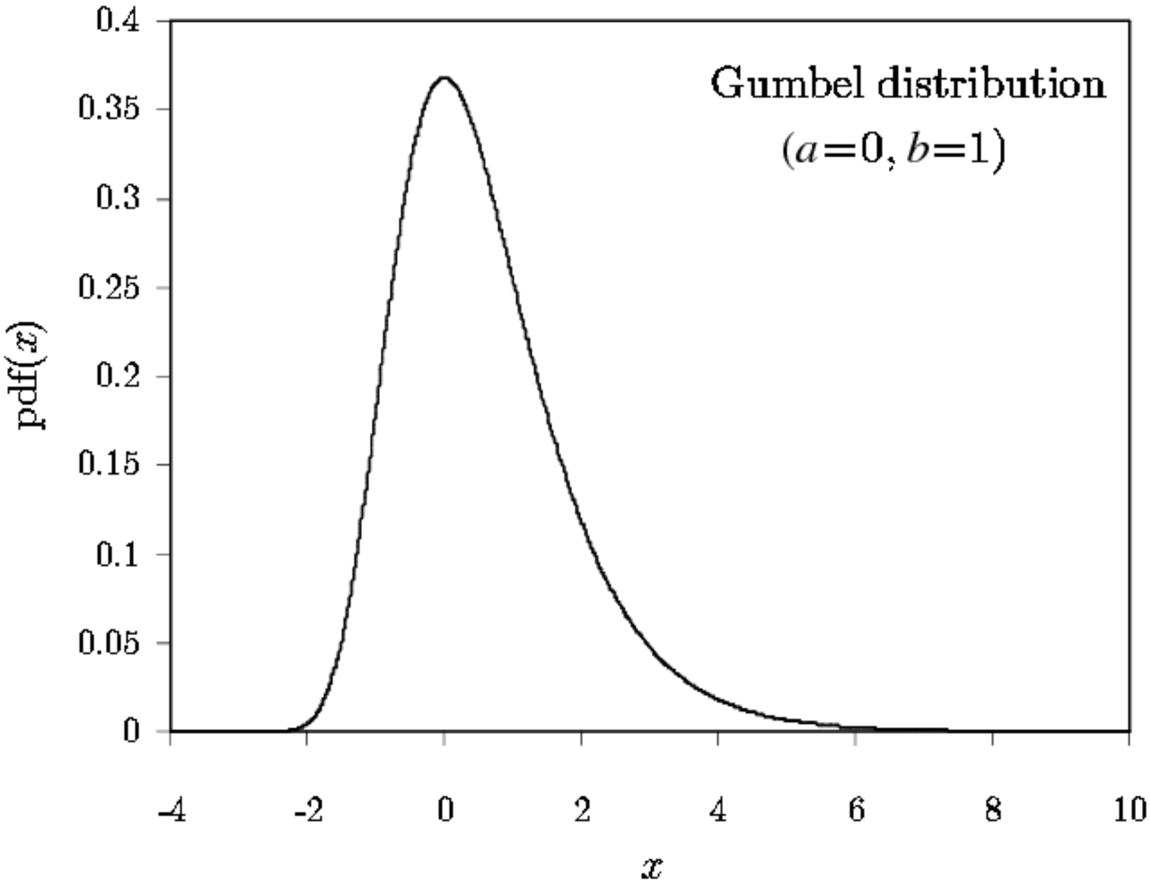
$$\frac{1}{1+10^{-(R_A-R_B)/D}}$$

for some constant D (often $D=400$) and the other pairwise probabilities are defined similarly. Now, look at what happens when we sample the "scores" for each player $S_A \sim \text{Gumbel}(R_A, D)$, $S_B \sim \text{Gumbel}(R_B, D)$, and $S_C \sim \text{Gumbel}(R_C, D)$. We want to calculate $P(S_A > S_B) = P(S_A - S_B > 0)$. Since S_A and S_B are drawn from Gumbel distributions with the same scale parameter, $S_A - S_B \sim \text{Logistic}(R_A - R_B, D)$. Using the CDF of the (base 10) logistic distribution,

$$P(S_A - S_B > 0) = \frac{1}{1+10^{-(R_A-R_B)/D}}$$

which is the same win probability that Elo predicted. This scales to any number of players $N \geq 2$.

In conclusion, we can simulate multi-player outcomes that are consistent with the pairwise win probabilities. My next goal is to calculate the "distribution" of where each player will finish. For example, Player A will finish in first place with probability p_{1A}, second place with probability p_{2A}, and so on. It's simple to do with a simulation, but if anyone has a way to derive those probabilities analytically I would be interested.



lity density function of the standard Gumbel distribution.

Note std gumbel is NOT symmetrical.

7.0 Unequal games

Teams play unequal games. Some play 12, some play up to 15. One plays 120. Not a problem.

OK, so now we have our mathematical method to computer rate teams. We can use this method each week as we get more trial outcomes (games), and therefore assign a more correct winning probability to each team. Great! But this does not answer the question that needs to be answered, specifically what 4 teams will make the top 4 spots at the end of the season? In order to do that we have to recognize the committee is composed of humans and humans have biases. The question then becomes "how can we measure the biases of the committee?". Once we accurately know their biases we can add those back into the computer rankings – and the computer will spit out the predicted committee rankings for the subsequent week.

8.0 model applied to specific real-life scenarios

Texas – Oklahoma – Texas Tech in 2008
One of the oldest debates in college football is A>B>C>A. How do you decide?

SECTION 2

USING BIAS TO PREDICT COMMITTEE RANKINGS

9.0 Committee bias and committee predictions

The method will assign a bias to a team based on the difference from the committee's ranking and the computers ranking, using the rating difference between those two rankings.

Using an example to illustrate, say the computer ranks Penn State #1 with a rating of .955 and the computer ranks Alabama #2 with a rating of .920. That same week if the committee ranks Alabama #1 and Penn State #2, then Alabama will be assigned a bias of .955-.920 (=+.035) and Penn State will be assigned a bias of .920-.955 (=-.035).

For a different example, assume the committee rankings are #1 Alabama, #2 Georgia, and #3 Penn State, while the computer rankings and ratings are #1 Penn State=0.955, #2 Alabama=0.920, and #3 Georgia=0.900 then the bias for each team would be:

- Alabama = 0.955-0.920 = +0.035
- Georgia = 0.920-0.900 = +0.020
- Penn State = 0.900 – 0.955 = -0.055

If in both the computer poll and committee poll Georgia is ranked #2, the bias is 0, regardless of what the computer rating of Georgia is.

The bias is then averaged over the course of a year and used to calculate the predicted committee rankings after games are played on Saturday, but before committee rankings are released on Tuesday. Bias is not carried over from year to year.

I have decided to move to this method because there is such a difference in computer rating between an undefeated team and a 1 loss top team. Because the committee only assigns rankings and not numerical ratings, it is very hard to capture this difference year to year. For example, there is a strong difference between 2016 Alabama at week 14, with a 13-0 record and ranked #1 and 2015 Michigan State at week 14, with a 12-1 record and also ranked #1. This method is the easiest way to not let that difference unduly influence bias.

To predict the playoff committee poll on Tuesday, the numerical rating from the computer is added to the bias for a team, and those ranks predict the committee rankings. As long as the bias remains about the same this method will track the committee poll with excellent results.

SECTION 3

USING COMPUTER RATINGS TO PREDICT FUTURE GAME OUTCOMES

10.0 Using ratings to compute future probabilities

Ratings applied to probability theory
Once we believe we have a true mathematical rating for 2 given teams, is there a way to determine the probability of team A beating team B? Is there a mathematical method to determine the expected point spread between team A and team B?

We notice the full ratings $\vec{r}$ are a normally distributed random variable with mean $\mu = 0.5$ and standard deviation $\sigma = 0.25 \pm \sim.1$

We move to exponent probability models, like are used in competitive chess. The equation takes the form of

$$P(A) = \frac{1}{1 + B^{[(r_A - r_B)/D]}}$$

using B base and D divisor. In chess, the values used are $B = 10$ and $D = 400$. The value of $D = 400$ leads to spread in ratings – A great rating is about 2,200 and a bad rating is about 600. This is a range of 1,600 or 4 blocks of 400 points each. PlayoffPredictor.com ratings go from about 0 to about 1, so if we wanted to segment that into 4 blocks we would use $D = 0.25$ and then select a B. However, B and D are related. Look at the following table of equivalent B and D pairs:

equivalent pairs					
base		10	100	1000	10000
divisor		1	2	3	4
base		10	100	1000	10000
divisor		0.5	1	1.5	2
base	e	10	100	1000	10000
divisor	0.1447705	0.3	0.666666	1	1.333333

Taking advantage of that we simply set $D = 1$ and solve for B.

1000 and 1 give the same result as 10000 and 1.3333, or 100 and .6666. In the same way 1000 and 3 give the same result as 100 and 2, or 10 and 1. This is the result of log math.
So we use a base of 1000 and a divisor of 1. The divisor of 1 make sense — take that out of the equation. Then, what base should you use? Empirically 1000 fits well.

At a base of 1000 and divisor of 1, a team with a rating +.16 more than an opponent will have a 75% chance of winning. So in this sense +.16 corresponds to 200 points in chess Elo.

In my rating the teams will be normally distributed with a mean at 0.5 and a standard deviation around 0.25. Meaning teams that are separated by 1 standard deviation, the better team has a 85% chance at success. For example, the following teams are all about one sigma apart:

- #1 Georgia (~1)
- #20 BYU (~.75)
- #70 Illinois (~.5)
- #108 Tulane (~.25)
- #129 1AA (FCS) (~0)

So Georgia has a 85% chance of beating BYU, BYU has a 85% chance of beating Illinois, Illinois has an 85% chance of beating Tulane, and Tulane has a 85% chance of beating a FCS school. Is that right? It does all correspond well to Vegas projected lines.

Keeping with the logic, Georgia would have a 97% chance of beating Illinois [1/(1+1000^(-.5))]. BYU would have a 97% chance of beating Tulane. Are those right? Again, they correlate well with Vegas.

We can turn those probabilities to point spreads using the following table. These numbers are derived from a fitting of the data. The original article is from boydsbets.com:

Point Spread	Road win probability
undef	0.00%
-40	0.10%
-39	0.30%
-38	0.60%
-37	1.00%
-36	1.20%
-35	1.40%
-34	1.60%
-33	1.80%
-32	2.00%
-31	2.20%
-30	2.50%
-29	2.70%
-28	2.90%
-27	3.20%
-26	3.40%
-25	3.60%
-24	3.80%
-23	4.50%
-22	5.10%
-21	5.30%
-20	6.00%
-19	6.30%
-18	8.60%
-17	11.00%
-16	11.50%
-15	12.60%
-14.5	14.90%
-14	16.50%
-13.5	17.00%
-13	17.40%
-12.5	18.40%
-12	19.40%
-11.5	20.10%
-11	20.80%
-10.5	22.60%
-10	24.50%
-9.5	25.00%
-9	25.40%
-8.5	26.20%
-8	27.00%
-7.5	29.70%
-7	32.30%
-6.5	33.60%
-6	34.90%
-5.5	35.90%
-5	36.90%
-4.5	38.10%
-4	39.40%
-3.5	42.60%
-3	45.80%
-2.5	46.60%
-2	47.50%
-1.5	48.80%
-1	49.50%
-0.5	*49.99%*
0	*50.01%*

Point Spread	Road win probability
0.5	50.50%
1	51.20%
1.5	52.50%
2	53.40%
2.5	54.30%
3	57.40%
3.5	60.60%
4	61.90%
4.5	63.10%
5	64.10%
5.5	65.10%
6	66.40%
6.5	67.70%
7	70.30%
7.5	73.00%
8	73.80%
8.5	74.60%
9	75.10%
9.5	75.50%
10	77.40%
10.5	79.30%
11	79.90%
11.5	80.60%
12	81.60%
12.5	82.60%
13	83.00%
13.5	83.50%
14	85.10%
14.5	86.80%
15	87.40%
16	88.60%
17	91.40%
18	93.70%
19	95.00%
20	96.20%
21	96.60%
22	97.00%
23	97.30%
24	98.00%
25	98.40%
26	98.70%
27	98.90%
28	99.50%
29	99.80%
30	100.00%

The range is from a -40-point home favorite to +30-point road favorite

It is Intentional to have more probability along multiples of 7 and 7x+3
This takes home field advantage into account in point spread, not favorite. For example, a 95% victory probability for the road team is a +19 point road favorite, where a 95% victory probability for a home team is a -23 point home favorite. The numbers are not symmetric.

The use of these values withing a range is from start to end inclusive (,] (excel match mode 1, search mode 1), an exact search or next larger value, starting at beginning. For example P(Road Team Wins) = 2.0% selects -32 point spread, just like P(Road Team Wins) = 1.99% also selects -32, whereas P(Road Team Wins) = 2.01% selects -31 point spread.

To test this method, we can look at a small sample from the 2022 beginning of bowl season. The below table lists the first 10 bowl games played, with the computer predicted winner and line, the Vegas line, and the actual score differential.

base used->	1000
home adv->	0
Multiplier ->	1

											Computer				Vegas				
id	home	away	my rating home	my rating away	percentage away	my predicted point spread	vegas point spread	Pred winner = vegas winner?	Bias - prediction to Vegas: average of (prediction - actual)	Actual	Straight up win percentage (prediction)	Spread win percentage (prediction)	average of (Absolute value of (prediction - actual))	average of (squares of (absolute error))	Straight up win percentage (Vegas)	Spread win percentage (Vegas) [0-0-N by definition]	average of (Absolute value of (prediction - actual))	average of (squares of (absolute error))	Improvement over Vegas
						-0.79	-1.57	0.80	0.78		0.70	0.75	9.60	153	0.80	0.50	10.20	143	0.60
											(7 - 3)	(7 - 2 - 1)			(8 - 2)	(0 - 0 - 10)			
										Over Expected ->	*-0.100*	*0.250*	*-0.60*	*10*					
401441993	UAB	Miami (OH)	0.42	0.33	35.68%	-5.5	-11	1	5.5	-4	1	1	1.5	2	1	0.5	7	49	+5.5
401441994	Troy	UTSA	0.77	0.74	43.43%	-3	-2	1	-1	-6	1	1	3	9	1	0.5	4	16	+1.0
401441995	Louisville	Cincinnati	0.66	0.66	49.34%	-1	-2.5	1	1.5	-17	1	0	16	256	1	0.5	14.5	210	-1.5
401441997	BYU	SMU	0.53	0.57	55.56%	3	4.5	1	-1.5	-1	0	1	4	16	0	0.5	5.5	30	+1.5
401442038	Fresno State	Washington State	0.60	0.60	50.53%	1	-5.5	0	6.5	-23	0	0	24	576	1	0.5	17.5	306	-6.5
401442041	Southern Miss	Rice	0.44	0.30	28.15%	-7.5	-7	1	-0.5	-14	1	0.5	6.5	42	1	0.5	7	49	+0.5
401441999	Oregon State	Florida	0.79	0.61	23.11%	-10	-7.5	1	-2.5	-27	1	1	17	289	1	0.5	19.5	380	+2.5
401442001	Boise State	North Texas	0.59	0.42	23.71%	-10	-12	1	2	-3	1	1	7	49	1	0.5	9	81	+2.0
401442003	Uconn	Marshall	0.38	0.59	81.29%	12	11.5	1	0.5	12	1	1	0	0	1	0.5	0.5	0	+0.5
401442004	San Jose State	Eastern Michigan	0.46	0.42	43.04%	-3	-3.5	1	0.5	14	0	1	17	289	0	0.5	17.5	306	+0.5

The model has an average predicted point spread of -0.79, close to Vegas' -1.57. Furthermore, with regard to actual results, this sample has a spread win percentage of 0.75, and mean average error (MAE) of 9.6. Vegas has a MAE of 10.2 for these games. The mean squared error is slightly higher than Vegas because of the Fresno State – Washington State game – actual was -23 where the model prediction was +1. The value of 24^2 blows up the total error significantly more than the Vegas error of 17.5^2.

Conclusions