PlayoffPredictor.com

an

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

By Neville Aga @playoffPredict
September 2023

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

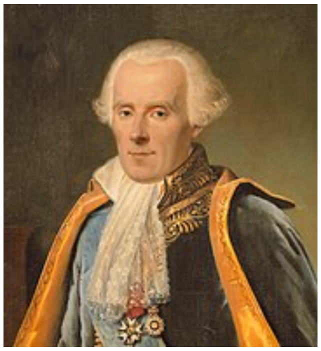
$$P(sun\ will\ rise\ tomorrow) = \frac{d+1}{d+2}$$

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

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

The sum of wins and losses equals the total games played, $n_{tot} = n_w + n_l$

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

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)

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

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

$$r_i = \frac{1 + n_{w,i}^{eff}}{2 + n_{tot,i}} \quad (4)$$

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.

[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.

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:

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

where

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

Substituting in the equality for n_w^{eff},

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

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

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

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. 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:

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

$n_{w,i}^{eff}$ is defined as it was in eq . 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 greater than average point differential by team i .
m_{ij} is defined to be zero for a win by team i with an average scoring differential.
m_{ij} is a negative number for a win by less than average point differential by team i .
And by necessity $m_{ij} = -m_{ji}$[4].

In a win by team i, m_{ij} as a number close to +1 means that team i blew out team j, a number close to 0 means team i and team j played a standard game (that i won), and a number close to -1 means team i barely squeaked by team j.
If team i lost the game, then m_{ij} as a number close to -1 means that team i got blown out by team j, a number close to 0 means team i and team j played a standard game (that j won), and a number close to +1 means team i barely lost to team j.

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 rating) is, but we are told "you landed left (won the game), but very close to the marker (m_{ij} is very close

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

to -1)", or "you landed left(won the game) by an average distance from that marker (m_{ij} is about zero)", or "you landed left (still won the game), far away from the marker (m_{ij} is very close to +1)". This is the conceptual value of m, our margin of victory.

Note inside m_i^{eff} that absolute score values are not considered – that is there is no $m_{tot,i}$ for the team independent from being multiplied by a team rating[5]. The only margins are multiplied by the team played or by the team itself (for the $m_{tot,i}$ in the denominator of equation 8). 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 scoring margins that deviate from median scoring margins, so if a team wins all their games by blowouts this will be a large positive number. Interestingly, if a team loses all their games (0-12) by squeakers $\sum_{j=1}^{n_{tot,i}} r_j^i \cdot m_{ij}$ will also be a large **positive** number. If they win half and lost half all by average margins, this will be about 0, and if they lose all their games by blowouts 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[6]. 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 8 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:

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

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

---

[5] Stated another way, the column vector b remains unchanged from the Colley definition.

[6] Imagine a 6-6 Mississippi State team. They count among their wins a 35-point victory against Alabama ($r_{Alabama} = 0.9$). They also count among their losses a 35-point loss to Vanderbilt ($r_{Vanderbilt} = 0.3$). The net margin for these 2 games will be +(0.9)(0.25)+(0.3)(-0.25) = +0.15, which is a net positive to Mississippi State.

A value of $m_{AB} = 0$ means that team A and B finished with no margin of victory (the median score differential in college football is 15 points) 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. A value of -1 for m_{AB} means that team A beat team B by the tiniest of margins, like a 1-point victory[7]. 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. 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

$$r_A \;=\; 0.75; \quad r_B \;=\; 0.25 \tag{9}$$

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

In matrix form, the equations become:

$$\vec{r} = A^{-1}\vec{b}, \tag{10}$$

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 $\vec{b}$, is the same column-vector as defined before.

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,

[7] Or better yet, like a victory by default, if such a thing were to ever happen.

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

The matrix A is the **playoffPredictor.com Matrix**. It extends the Colley Matrix with information on margin of victory. *Solving equations (10)–(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:

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

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

Which recovers our $r_A = 0.75$ and $r_B = 0.25$ that we obtained in equation (9)

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)
$\sum b = count_{teams}$,
$\sum A^{-1} = 0.5 * count_{teams}$, and
$\sum A = 2 * count_{teams}$

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

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

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

(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 #2: 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.

2.3 Practical value of m

The correct way to choose m for any given gave is to scale margin-of-victory from the point domain into the time domain. Since our equations will effectively give a winning team A credit for a little extra victory over team B, we want to match the extra quantity to the quantity of the win. That sounds odd because winning is a binary even (happens or does not), but we strive to stretch that definition here. If the game is effectively over at the first minute after halftime then m should be about +0.5, since that is the amount of game left. If we simply choose the range for a close game to match, then m=-0.5 will indicate a game that came down to the final play. This gives us a working range well within the +1 (infinitely better) and -1 (same) theoretical m range.

Using inputs as a final score though, there is no elegant truth to the correct value of m for any given single game. A game that ends with a 1 point win can be decided much earlier than a game that ends with a 5 point victory, there simply is no way to infer that without more information than the final score. Nevertheless, since at this time the only input we have is the final score we can propose some common-sense mappings. The model-maker can choose to scale it piecewise linearly such that a 1-point victory is worth $m = -0.5$, a 15-point victory[8] (15 being the median margin of victory in college football) is worth $m = 0$, and an 80-point victory is worth $m = 0.5$, 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 7 points and 8 points – the former can be tied on one score at the end of the game, where the later needs a score and a 2-point conversion to tie.

The simplest option for the model maker is tiered Margin of Victory:

$$m = \begin{cases} -0.25, & for\ MoV\ of\ 1-7\ points \\ 0, & for\ MoV\ of\ 8-24\ points \\ 0.25, & for\ MoV\ of\ 25+\ points \end{cases}$$

This corresponds to thinking about wins as close wins, solid wins, and blowouts. This is closer to how humans view results – pollsters tend not to differentiate much between a 10-point win and a 20-point win. They are both solid wins. The advantage here is that this keeps the method simple, and our goal is simplicity for a "football truth". Wins are tiered into sub 1 score, 2-3 score wins, and 4+ score wins. Why this breakdown? Because in a given season roughly 1/3 of all games fall into each tier. In the 2021 season the specific breakdown was:

tier	Margin of Victory	Number of games
close wins	1-7	281
solid wins	8-24	332
blowouts	25+	267

(n=880 total games).

[8] A 15-point victory is expressly defined as $m = 0$, so the equation $\vec{r} = A^{-1}\vec{b}$, expressly simplifies to $\boldsymbol{\vec{r}} = \boldsymbol{C^{-1}\vec{b}}$

Can this tiering be improved upon? Probably. You could have more tiers as on the surface a 24-point victory seems a lot more solid than an 8-point victory. However, the value of these tiers is simplicity. It is easy to understand the idea of solid wins and blowouts, and simple to go in one-quarter increments.

Because the number of games in these 3 categories is all similar and we use a negative value for m in the close wins that is of equal magnitude to the positive value for m in the blowouts we arrive at an important result, namely $\frac{1}{3}(-.25) + \frac{1}{3}(0) + \frac{1}{3}(+.25) = 0$. That means that we will not have inflation of rating when we add in margin-of-victory. If instead we choose

$$m = \begin{cases} 0, & for\ MoV\ of\ 1-7\ points \\ 0.25, & for\ MoV\ of\ 8-24\ points \\ 0.50, & for\ MoV\ of\ 25+\ points \end{cases}$$

we would get similar *rankings* to our $m \in (-0.25\,, 0\,, +.25)$ model but very different *ratings*. That is because we will have rating inflation for the best teams and rating deflation for the worst – at the end of a season the ratings would be in the range of $-0.5 < r < +1.5$, instead of $0 < r < +1$. Keeping the ratings roughly between 0 and 1 allows for direct comparison to the Colley Matrix and easy comparisons over time to different teams from different seasons.

For our model we will use the following tiered margin-of-victory weights:

$$m = \begin{cases} -0.2, & for\ MoV\ of\ 1-2\ points \\ 0, & for\ MoV\ of\ 3-24\ points \\ 0.2, & for\ MoV\ of\ 25-34\ points \\ 0.3, & for\ MoV\ of\ 35+\ points \end{cases}$$

> ***The playoffPredictor.com computer model uses the option for tiered margin-of-victory with weights of $\boldsymbol{m} \in (-\mathbf{0.2}\,, \mathbf{0}\,, +\mathbf{0.2}\,, +\mathbf{0.3})$ for MoV $\in (\mathbf{1-2}, \mathbf{3-24}\,, \mathbf{25-34},\ \mathbf{35+})$ margin.***

2.4 Margin-of-Victory range

We have already seen how an $m = +1$ impacts a 2-team league. How does an $m = -1$ impact the same league? Recall that as α was increased from 0 to 0.5 the ratings increased from $r_A = 0.625;\ \ r_B = 0.375$ to $r_A = 0.75;\ \ r_B = 0.25$, now as the product of $m\alpha$ goes from 0 to -0.5 we find that $r_A = 0.5;\ \ r_B = 0.5$

15-point victory		
(m = 0, no MoV in this win)		
	# matches played	
teams	1	
A	0.625	
B	0.375	

1 point victory (m =-.25), α = 0.5

	# matches played	
teams	1	
A	0.594	
B	0.406	

victory with MoV such that (m= -0.5), α = 0.5

	# matches played	
teams	1	
A	0.563	
B	0.437	

victory with MoV such that (m= -1.0), α = 0.5

teams	# matches played	
	1	
A	0.500	
B	0.500	

We come to a conclusion that is the anthesis of the $m = +1$ scenario – namely a victory where $m = -1$ does not help the rating of A over B at all. It is like a $m = -1$ victory is a victory by default, the other team missed the bus. Note that this is not the same as the 2 teams never having played. Examine the matrix before any games are played:

$$\begin{matrix} r_A \\ r_B \end{matrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{matrix} 0.5 \\ 0.5 \end{matrix}$$

And examine the matrix after A defeats B with $m_{AB} = -1$

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

The final results may be all the same 0.5 rating, but the matrix in the middle is vastly different. This will have the effect of drawing the 2 teams closer to each other, a very different result when combined with all the other games in the season.

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 15 points and B beats C, also by 15 points. Our equations become:

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

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

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

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

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

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

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

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. Using tiered m our model gives a 21-point victory of $m_{AB} = m_{BC} = 0$ and a 41-point victory gives $m_{AC} = 0.25$:[9]

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

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

[9] And of course, $m_{CA} = -0.25$

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

$$[2 + 2 - 0.5(.25 - 0)]\, r_A = 1 + \left[\frac{(2-0)}{2} + r_B + r_C\right] + 0.5[\, r_B \cdot (0) + r_C \cdot (0.25)]$$

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

$$[2 + 2 - 0.5(0 - .25)]\, r_C = 1 + \left[\frac{(0-2)}{2} + r_A + r_B\right] + 0.5[\, r_A \cdot (-0.25) + r_B \cdot (0)]$$

$$[3.875]\, r_A = [2 + r_B + r_C] + [+0.125 r_C]$$
$$[4]\, r_B = [1 + r_A + r_C] + [0]$$
$$[4.125]\, r_C = [0 + r_A + r_B] + [-0.125 r_A]$$

$$3.875\, r_A - r_B - 1.125 r_C = 2$$
$$-r_A + 4\, r_B - r_C = 1$$
$$-0.875 r_A - r_B + 4.125\, r_C = 0$$

$$r_A = \frac{r_B + 1.125 r_C + 2}{3.875}$$

$$r_B = \frac{r_A + r_C + 1}{4}$$

$$r_C = \frac{0.875 r_A + r_B + 0}{4.125}$$

The values $r_A = 0.725$, $r_B = 0.5$, $r_C = 0.275$ solve these equations. This is an intuitive result. A is better and C is worse as a reward (punishment) for the blowout win as compared to when they were all standard wins.

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 15-point victories we saw $r_A = 0.7$, $r_B = 0.5$, $r_C = 0.3$. In the case of the 41-point victory for A over C and 21-point victories A-B & B-C, we see $r_A = 0.725$, $r_B = 0.5$, $r_C = 0.275$. (note that both 15 points and 21 points map to an $m = 0$)

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 a 21/21/41-point victories or 15/15/15-point victories, if they all scale together.

It is interesting to note that while 21/21/41 and 15/15/15 are mathematically consistent results, 1/1/1 and 41/41/41 are not. For 1/1/1 B's rating is just over 0.5 and for 41/41/41 B's rating is just under 0.5.

The model does produce interesting results if there are inconsistencies in the victories. For example, if A barely beats B (1-point), 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 (41-points). Then, to satisfy the equations, you arrive at $r_A = 0.65$, $r_B = 0.55$, $r_C = 0.30$. Notice that A does not get the 'credit' for the B destruction of C. Because B>>C, B's rating is increased by .05 at the expense of C. In a Z score context, A is +1.02, where B is +.35, and C is -1.37. See the chart below for all the permutations of 3 teams considering margin of victory of "barely beats" ($m = -0.25$), "beats" ($m = 0$), and "destroys" ($m = 0.25$).

3-team league results
Using win-margin perspective
All cases A>B>C & A>C

									rA	rB	rC	σA	σB	σC
A	destroys	B	B	barely beats	C	A	barely beats	C	0.760	0.440	0.300	+1.35	-0.31	-1.04
A	destroys	B	B	barely beats	C	A	beats	C	0.788	0.439	0.273	+1.34	-0.29	-1.06
A	destroys	B	B	barely beats	C	A	destroys	C	0.816	0.437	0.247	+1.34	-0.26	-1.07
A	destroys	B	B	beats	C	A	barely beats	C	0.761	0.458	0.282	+1.32	-0.21	-1.10
A	destroys	B	B	beats	C	A	beats	C	0.788	0.456	0.256	+1.31	-0.20	-1.11
A	destroys	B	B	beats	C	A	destroys	C	0.816	0.454	0.231	+1.31	-0.19	-1.12
A	beats	B	B	barely beats	C	A	barely beats	C	0.730	0.470	0.300	+1.30	-0.17	-1.13
A	beats	B	B	barely beats	C	A	beats	C	0.756	0.469	0.274	+1.29	-0.15	-1.14
A	beats	B	B	barely beats	C	A	destroys	C	0.783	0.469	0.248	+1.29	-0.14	-1.15
A	destroys	B	B	destroys	C	A	beats	C	0.789	0.473	0.239	+1.28	-0.12	-1.16
A	destroys	B	B	destroys	C	A	destroys	C	0.816	0.470	0.214	+1.281	-0.122	-1.159
A	destroys	B	B	destroys	C	A	barely beats	C	0.762	0.475	0.263	+1.28	-0.12	-1.16
A	beats	B	B	beats	C	A	barely beats	C	0.730	0.489	0.281	+1.25	-0.06	-1.19
A	beats	B	B	beats	C	A	beats	C	0.756	0.487	0.256	+1.25	-0.06	-1.19
A	beats	B	B	beats	C	A	destroys	C	0.782	0.486	0.231	+1.25	-0.06	-1.19
A	barely beats	B	B	barely beats	C	A	barely beats	C	0.700	0.500	0.300	+1.22	-0.00	-1.22
A	barely beats	B	B	barely beats	C	A	beats	C	0.725	0.500	0.275	+1.22	-0.00	-1.22
A	barely beats	B	B	barely beats	C	A	destroys	C	0.750	0.500	0.250	+1.22	0	-1.22
A	beats	B	B	destroys	C	A	destroys	C	0.782	0.504	0.214	+1.22	+0.02	-1.23
A	beats	B	B	destroys	C	A	beats	C	0.756	0.506	0.238	+1.21	+0.03	-1.24
A	beats	B	B	destroys	C	A	barely beats	C	0.731	0.507	0.262	+1.20	+0.04	-1.24
A	barely beats	B	B	beats	C	A	destroys	C	0.749	0.519	0.232	+1.18	+0.09	-1.27
A	barely beats	B	B	beats	C	A	beats	C	0.725	0.519	0.256	+1.17	+0.10	-1.27
A	barely beats	B	B	beats	C	A	barely beats	C	0.700	0.520	0.280	+1.16	+0.12	-1.28
A	barely beats	B	B	destroys	C	A	destroys	C	0.748	0.538	0.214	+1.13	+0.17	-1.30
A	barely beats	B	B	destroys	C	A	beats	C	0.724	0.539	0.237	+1.12	+0.19	-1.31
A	barely beats	B	B	destroys	C	A	barely beats	C	0.700	0.540	0.260	+1.10	+0.22	-1.32

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 beats C, and C barely beats A. In that case you have $r_A = 0.551$, but a standard deviation of +1.41, which is the greatest standard deviation in the table. Notice the rating $r_A = 0.552$ is higher if A destroys B, B destroys C, and C barely beats A, but the standard deviation is lower at +1.28.

3-team league results
Using win-margin perspective
All cases A>B>C>A

									rA	rB	rC	σA	σB	σC
A	destroys	B	B	beats	C	C	barely beats	A	0.551	0.473	0.476	+1.41	-0.76	-0.65
A	destroys	B	B	destroys	C	C	beats	A	0.526	0.498	0.476	+1.28	-0.12	-1.16
A	destroys	B	B	destroys	C	C	barely beats	A	0.552	0.495	0.453	+1.28	-0.12	-1.16
A	destroys	B	B	barely beats	C	C	barely beats	A	0.550	0.450	0.500	+1.22	-1.22	0
A	destroys	B	B	beats	C	C	beats	A	0.524	0.474	0.501	+1.19	-1.25	+0.06
A	destroys	B	B	barely beats	C	C	beats	A	0.522	0.451	0.526	+0.65	-1.41	+0.76
A	destroys	B	B	destroys	C	C	destroys	A	0.500	0.500	0.500	-	-	-
A	destroys	B	B	barely beats	C	C	destroys	A	0.495	0.453	0.552	-0.12	-1.16	+1.28
A	destroys	B	B	beats	C	C	destroys	A	0.498	0.476	0.526	-0.12	-1.16	+1.28

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.

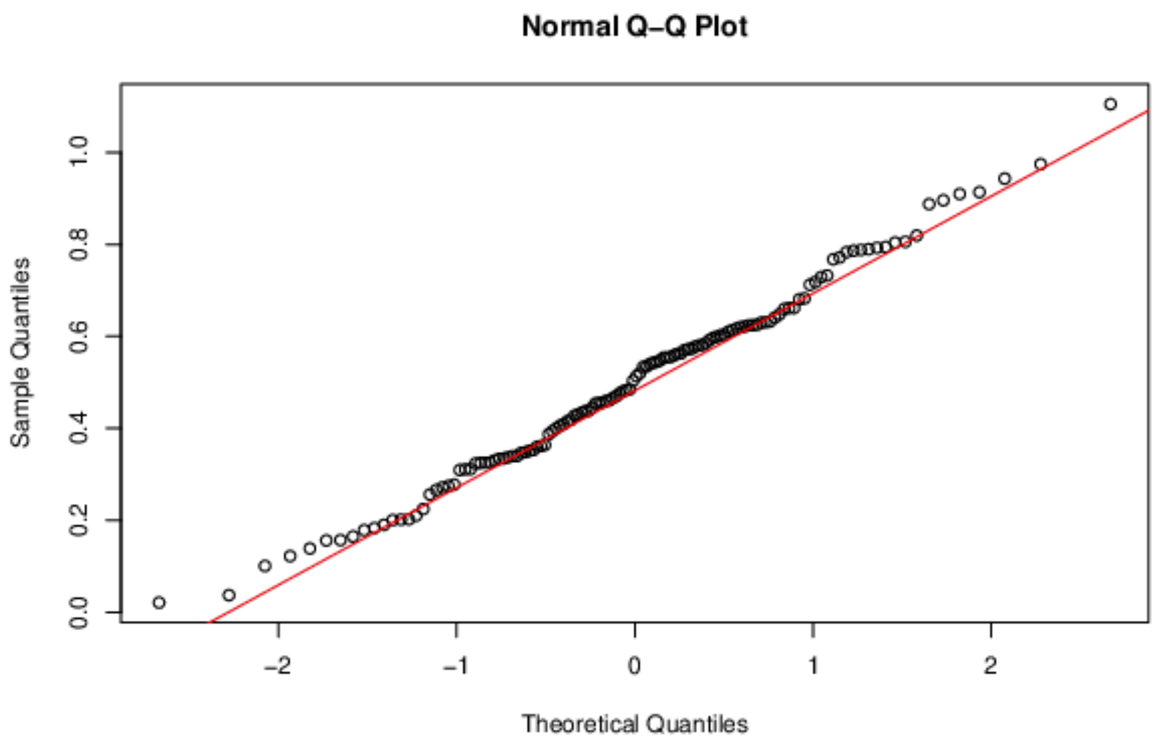
In the final AP poll before the conference championship week eight more voters decided to elevate Texas as the best group of that trio (over Oklahoma), while the coaches' poll had Oklahoma on top. Highlighted in blue in table X is the Oklahoma-Texas Tech-Texas scenario, with Oklahoma as team A, Texas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

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 top 10 team, the rating of the 12-0 team increases approximately 1.5X how its rating would increase for the same m point victory against the bottom 10 team (returning the top 10 team victory to 1-point). You are rewarded more for larger MoV victories against top teams, which is intuitive.

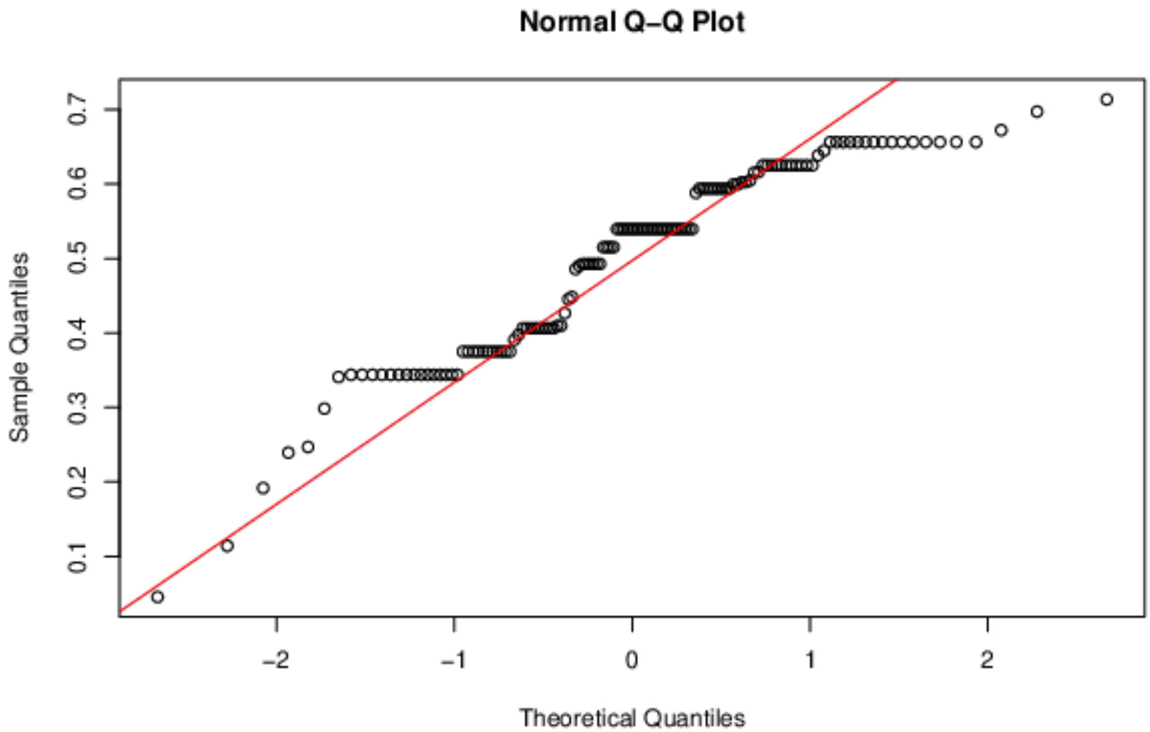
4.0 Normal Distribution of the ratings

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 final $\vec{r}$ for all teams in 2022 with the actual 2022 $\vec{r}$ plotted against theoretical:



The resulting plot line gets more linear as more weeks go on. Here is the Q-Q plot after just 1 week of games have concluded (2022).



The dot on the very bottom left is the group of FCS schools with a record of 2-39. The bubble on the top right is Alabama with a record of 1-0, with a $m = +.25$ victory over Utah State. Utah State at this time has a 0.4902 rating, with a 1-1 record, having played and won a week 0 game to UConn. The amazing thing here is this model ranks Alabama #1 after just 1 week of data with *no preseason or prior year information at all* even though 21 teams won in week 1 by the same m margin (25+ points) against FCS schools.

The ratings confirm a normal distribution after enough games are sampled.

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

5.1 expanded Colley 5-team analysis

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

A	10	C	0
A	0	D	10
A	0	E	10
B	0	C	10
B	10	E	0
C	10	D	0
C	0	E	10

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 10-point wins by the score of 10-0 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 10-point loss to team B.

A	10	C	0
A	0	D	10
A	0	E	40
B	0	C	10
B	10	E	0
C	10	D	0
C	0	E	10

Final standings	final rating
E	0.604
B	0.526
C	0.500
D	0.474
A	0.396

Note that A gets significantly stronger, moving up +.017, 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 (the point they both get tied is at a .558 rating).

5.2 2008 ACC 5-team analysis

Let's examine another 5-team scenario, the 2008 ACC subset of Miami, VT, UNC, UVA, and Duke. All teams played each other exactly 1 time with the following results:

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

To start, we need a metric to compare against. We will use the term "mean absolute ratio" denoted by η, as first proposed by Wes Colley in his 2002 paper.

$$\eta = \exp\left((\frac{1}{25} \sum_{i=1}^{25} |\ln j_C(team_i) - \ln i| \right)$$

η is the ratio of the playoffPredictor ranking to the poll ranking, or vice versa, so that the larger of the two always in the numerator, (specifically η is the exponent of the mean of the absolute values of the logs of the ratios), so η = 1.25 means the rankings would differ by typically one place at #4, and 5 places at #20.

Now that we have our metric in place, we backtest with 2022 data. Each week the games for the previous weeks are input and ratings generated. Those ratings are compared with η to the polls (AP for weeks 1-9, cfp committee for weeks 10-14, and AP for final rankings).

2022 playoffPredictor weekly η poll results

	m=0 (CM)	m=0.1	m=0.2	m=0.25	m=0.3	m=1
week 0 η	---					
week 1 η	2.93	2.71	2.70	2.73	2.76	2.95
week 2 η	3.13	2.95	2.83	2.78	2.68	2.73
week 3 η	2.81	2.72	2.56	2.44	2.41	2.40
week 4 η	2.11	2.04	2.03	2.00	1.99	2.64
week 5 η	1.79	1.76	1.73	1.72	1.76	2.42
week 6 η	1.52	1.43	1.42	1.42	1.37	2.00
week 7 η	1.63	1.40	1.36	1.32	1.32	1.84
week 8 η	1.50	1.44	1.37	1.39	1.38	1.70
week 9 η	1.36	1.30	1.30	1.36	1.37	2.00
week 10 η	1.21	1.24	1.28	1.30	1.33	2.06
week 11 η	1.19	1.16	1.19	1.27	1.29	1.77
week 12 η	1.18	1.18	1.21	1.22	1.26	1.95
week 13 η	1.17	1.20	1.21	1.25	1.28	1.90
week 14 η	1.18	1.20	1.24	1.30	1.31	1.83
end-of-season η	1.25	1.26	1.31	1.33	1.34	1.79

In green I have highlighted the best η for each week. Generally, η is best in early weeks with a weighting of around m=0.3, and best in late weeks with an m=0 (which is the Colley method, no margin-of-victory since $\alpha m = 0$ when $m = 0$). The best improvement on the m=0 weight is in week 7 where η is improved from 1.63 to 1.32 (for both m=0.25 and m=0.3) which is an improvement of 49%! Improvement in η is calculated as:

$$\eta_{improvement} = \frac{\eta_{m=0} - \eta_{m=X}}{\eta_{m=0} - 1}$$

Where $\eta_{m=X}$ means we are comparing the improvement from weights of $m = 0$ to $m = X$, which in the playoffPredictor.com method is 0.25

(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	Committee/Poll Week 5		Diff to Aga-elo mat	| ln A(i)−ln C(i) |	Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(5-0)	1	Ohio State	0.8414	+1.959		1	Alabama	1	0.693	1	Georgia	4	1.609
(5-0)	2	Alabama	0.8407	+1.955		2	Georgia	3	0.916	2	Michigan	9	1.705
(5-0)	3	Clemson	0.8276	+1.880		3	Ohio State	-2	1.099	3	TCU	13	1.674
(5-0)	4	Kansas	0.8075	+1.764		4	Michigan	7	1.012	4	Ohio State	-3	1.386
(5-0)	5	Georgia	0.8018	+1.732		5	Clemson	-2	0.511	5	Alabama	-3	0.916
(5-0)	6	Ole Miss	0.8006	+1.725		6	USC	2	0.288	6	Tennessee	4	0.511
(5-0)	7	Penn State	0.7956	+1.696		7	Oklahoma State	13	1.050	7	Clemson	-4	0.847
(5-0)	8	USC	0.7884	+1.655		8	Tennessee	2	0.223	8	Utah	28	1.504
(5-0)	9	UCLA	0.7826	+1.622		9	Ole Miss	-3	0.405	9	Kansas State	17	1.061
(4-0)	10	Tennessee	0.7776	+1.593	0.366	10	Penn State	-3	0.357	10	USC	-2	0.223
(5-0)	11	Michigan	0.7763	+1.586		11	Utah	25	1.186	11	Penn State	-4	0.452
(5-0)	12	Syracuse	0.7664	+1.529		12	Oregon	5	0.348	12	Washington	16	0.847
(4-1)	13	Mississippi State	0.7661	+1.527		13	Kentucky	16	0.802	13	Florida State	9	0.526
(4-1)	14	Wake Forest	0.7561	+1.470		14	NC State	10	0.539	14	Oregon State	43	1.404
(4-0)	15	James Madison	0.7480	+1.423		15	Wake Forest	-1	0.069	15	Oregon	2	0.125
(4-0)	16	TCU	0.7277	+1.307		16	BYU	7	0.363	16	Tulane	14	0.629
(4-1)	17	Oregon	0.7222	+1.275		17	TCU	-1	0.061	17	LSU	1	0.057
(4-1)	18	LSU	0.7145	+1.231		18	UCLA	-9	0.693	18	UCLA	-9	0.693
(5-0)	19	Coastal Carolina	0.7044	+1.173		19	Kansas	-15	1.558	19	South Carolin	52	1.318
(4-0)	20	Oklahoma State	0.7040	+1.171		20	Kansas State	6	0.262	20	Texas	19	0.668
(4-1)	21	Liberty	0.6945	+1.116		21	Washington	7	0.288	21	Notre Dame	25	0.784
(4-1)	22	Florida State	0.6926	+1.105		22	Syracuse	-10	0.606	22	Mississippi St	-9	0.526
(4-1)	23	BYU	0.6856	+1.065		23	Mississippi State	-10	0.571	23	NC State	1	0.043
(4-1)	24	NC State	0.6839	+1.055		24	Cincinnati	10	0.348	24	Troy	29	0.792
(4-1)	25	Memphis	0.6835	+1.053		25	LSU	-7	0.329	25	UTSA	43	1.001
(4-1)	26	Kansas State	0.6831	+1.051				calculated η →	1.791			calculated η →	2.345
(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			η (α)	η improvement			η (α)	η improvement	
(4-1)	31	Duke	0.6310	+0.752		η (0)	1.791	-		η (0)	2.345	-	
(4-1)	32	Washington State	0.6294	+0.742									

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	Committee/Poll Week 5		Diff to Aga-elo mat	| ln A(i)−ln C(i) |	Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(5-0)	1	Alabama	1.0089	+2.342		1	Alabama	0	0	1	Georgia	2	1.099
(5-0)	2	Ohio State	1.0043	+2.321		2	Georgia	1	0.405	2	Michigan	3	0.916
(5-0)	3	Georgia	0.9148	+1.909		3	Ohio State	-1	0.405	3	TCU	12	1.609
(5-0)	4	Clemson	0.8916	+1.802		4	Michigan	1	0.223	4	Ohio State	-2	0.693
(5-0)	5	Michigan	0.8842	+1.768		5	Clemson	-1	0.223	5	Alabama	-4	1.609
(5-0)	6	Ole Miss	0.8821	+1.758		6	USC	1	0.154	6	Tennessee	4	0.511
(5-0)	7	USC	0.8684	+1.696		7	Oklahoma State	11	0.944	7	Clemson	-3	0.560
(5-0)	8	Penn State	0.8681	+1.694		8	Tennessee	2	0.223	8	Utah	21	1.288
(5-0)	9	Kansas	0.8671	+1.690		9	Ole Miss	-3	0.405	9	Kansas State	11	0.799
(4-0)	10	Tennessee	0.8656	+1.682	0.660	10	Penn State	-2	0.223	10	USC	-3	0.357
(4-1)	11	Mississippi State	0.8402	+1.566		11	Utah	18	0.969	11	Penn State	-3	0.318
(5-0)	12	UCLA	0.8357	+1.545		12	Oregon	7	0.460	12	Washington	14	0.773
(4-0)	13	James Madison	0.8307	+1.522		13	Kentucky	19	0.901	13	Florida State	8	0.480
(5-0)	14	Syracuse	0.8296	+1.517		14	NC State	9	0.496	14	Oregon State	44	1.421
(4-0)	15	TCU	0.8029	+1.394		15	Wake Forest	1	0.065	15	Oregon	4	0.236
(4-1)	16	Wake Forest	0.7937	+1.352		16	BYU	12	0.560	16	Tulane	18	0.754
(4-1)	17	LSU	0.7871	+1.321		17	TCU	-2	0.125	17	LSU	0	0
(4-0)	18	Oklahoma State	0.7560	+1.178		18	UCLA	-6	0.405	18	UCLA	-6	0.405
(4-1)	19	Oregon	0.7368	+1.090		19	Kansas	-10	0.747	19	South Carolin	55	1.360
(4-1)	20	Kansas State	0.7281	+1.050		20	Kansas State	0	0	20	Texas	11	0.438
(4-1)	21	Florida State	0.7251	+1.036		21	Washington	5	0.214	21	Notre Dame	21	0.693
(4-1)	22	Maryland	0.7243	+1.032		22	Syracuse	-8	0.452	22	Mississippi St	-11	0.693
(4-1)	23	NC State	0.7222	+1.023		23	Mississippi State	-12	0.738	23	NC State	0	0
(5-0)	24	Coastal Carolina	0.7077	+0.956		24	Cincinnati	9	0.318	24	Troy	31	0.829
(4-1)	25	Liberty	0.7068	+0.952		25	LSU	-8	0.386	25	UTSA	40	0.956
(4-1)	26	Washington	0.6950	+0.898				calculated η →	1.494			calculated η →	2.121
(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			η (α)	η improvement			η (α)	η improvement	
(3-2)	31	Texas	0.6738	+0.800		η (0)	1.791	-		η (0)	2.345	-	
(4-1)	32	Kentucky	0.6715	+0.789		η (.5)	1.494	37%		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:

mean rating	0.500	Green cells are for user input	
std deviation	0.188	α	0
		yellow cells are calculated	
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

record	#	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) |	Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
1	Tennessee	3	1.386	1	Georgia	4	1.609
2	Ohio State	-1	0.693	2	Michigan	4	1.099
3	Georgia	2	0.511	3	TCU	-1	0.405
4	Clemson	-1	0.288	4	Ohio State	-3	1.386
5	Michigan	1	0.182	5	Alabama	2	0.336
6	Alabama	1	0.154	6	Tennessee	-2	0.405
7	TCU	-5	1.253	7	Clemson	-4	0.847
8	Oregon	1	0.118	8	Utah	9	0.754
9	USC	-1	0.118	9	Kansas State	3	0.288
10	LSU	5	0.405	10	USC	-2	0.223
11	Ole Miss	0	0	11	Penn State	7	0.492
12	UCLA	-2	0.182	12	Washington	23	1.070
13	Kansas State	-1	0.080	13	Florida State	18	0.869
14	Utah	3	0.194	14	Oregon State	10	0.539
15	Penn State	3	0.182	15	Oregon	-6	0.511
16	Illinois	3	0.172	16	Tulane	-3	0.208
17	North Carolina	-1	0.061	17	LSU	-2	0.125
18	Oklahoma State	3	0.154	18	UCLA	-8	0.588
19	Tulane	-6	0.379	19	South Carolin	27	0.884
20	Syracuse	-6	0.357	20	Texas	10	0.405
21	Wake Forest	2	0.091	21	Notre Dame	7	0.288
22	NC State	-2	0.095	22	Mississippi St	7	0.276
23	Oregon State	1	0.043	23	NC State	-3	0.140
24	Texas	6	0.223	24	Troy	2	0.080
25	UCF	9	0.307	25	UTSA	8	0.278
		calculated η →	1.357			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

mean rating	0.500	Green cells are for user input	
std deviation	0.234	α	0.5
		yellow cells are calculated	
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

record	#	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) |	Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
1	Tennessee	2	1.099	1	Georgia	1	0.693
2	Ohio State	-1	0.693	2	Michigan	2	0.693
3	Georgia	-1	0.405	3	TCU	3	0.693
4	Clemson	3	0.560	4	Ohio State	-3	1.386
5	Michigan	-1	0.223	5	Alabama	0	0
6	Alabama	-1	0.182	6	Tennessee	-3	0.693
7	TCU	-1	0.154	7	Clemson	0	0
8	Oregon	2	0.223	8	Utah	6	0.560
9	USC	-1	0.118	9	Kansas State	3	0.288
10	LSU	3	0.262	10	USC	-2	0.223
11	Ole Miss	0	0	11	Penn State	6	0.435
12	UCLA	-3	0.288	12	Washington	21	1.012
13	Kansas State	-1	0.080	13	Florida State	19	0.901
14	Utah	0	0	14	Oregon State	10	0.539
15	Penn State	2	0.125	15	Oregon	-5	0.405
16	Illinois	2	0.118	16	Tulane	-1	0.065
17	North Carolina	2	0.111	17	LSU	-4	0.268
18	Oklahoma State	2	0.105	18	UCLA	-9	0.693
19	Tulane	-4	0.236	19	South Carolin	31	0.968
20	Syracuse	-4	0.223	20	Texas	2	0.095
21	Wake Forest	2	0.091	21	Notre Dame	5	0.214
22	NC State	-1	0.047	22	Mississippi St	6	0.241
23	Oregon State	1	0.043	23	NC State	-2	0.091
24	Texas	-2	0.087	24	Troy	6	0.223
25	UCF	6	0.215	25	UTSA	12	0.392
		calculated η →	1.256			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

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

		team	rating	Z*-score	Z* range 1-10	Committee/Poll Week 13		Diff to Aga-elo mat	| ln A(i)−ln C(i) |	Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(12-0)	1	Georgia	0.9762	+2.440		1	Georgia	0	0	1	Georgia	0	0
(12-0)	2	TCU	0.9456	+2.284		2	Michigan	1	0.405	2	Michigan	1	0.405
(12-0)	3	Michigan	0.9196	+2.150		3	TCU	-1	0.405	3	TCU	-1	0.405
(11-1)	4	USC	0.8789	+1.942		4	USC	0	0	4	Ohio State	1	0.223
(11-1)	5	Ohio State	0.8692	+1.892		5	Ohio State	0	0	5	Alabama	2	0.336
(10-2)	6	Tennessee	0.8319	+1.701		6	Alabama	1	0.154	6	Tennessee	0	0
(10-2)	7	Alabama	0.8267	+1.674		7	Tennessee	-1	0.154	7	Clemson	2	0.251
(10-2)	8	Penn State	0.8149	+1.614		8	Penn State	0	0	8	Utah	9	0.754
(10-2)	9	Clemson	0.7941	+1.507		9	Clemson	0	0	9	Kansas State	1	0.105
(9-3)	10	Kansas State	0.7740	+1.404	1.036	10	Kansas State	0	0	10	USC	-6	0.916
(9-3)	11	Oregon	0.7635	+1.350		11	Utah	6	0.435	11	Penn State	-3	0.318
(10-2)	12	Washington	0.7610	+1.337		12	Washington	0	0	12	Washington	0	0
(9-3)	13	LSU	0.7532	+1.298		13	Florida State	3	0.208	13	Florida State	3	0.208
(8-4)	14	Texas	0.7428	+1.244		14	LSU	-1	0.074	14	Oregon State	1	0.069
(9-3)	15	Oregon State	0.7415	+1.238		15	Oregon State	0	0	15	Oregon	-4	0.310
(9-3)	16	Florida State	0.7394	+1.227		16	Oregon	-5	0.375	16	Tulane	2	0.118
(9-3)	17	Utah	0.7351	+1.205		17	UCLA	2	0.111	17	LSU	-4	0.268
(10-2)	18	Tulane	0.7307	+1.182		18	Tulane	0	0	18	UCLA	1	0.054
(9-3)	19	UCLA	0.7280	+1.169		19	South Carolina	9	0.388	19	South Carolin	9	0.388
(10-2)	20	Troy	0.7234	+1.145		20	Texas	-6	0.357	20	Texas	-6	0.357
(8-4)	21	Mississippi State	0.7027	+1.039		21	Notre Dame	1	0.047	21	Notre Dame	1	0.047
(8-4)	22	Notre Dame	0.6969	+1.009		22	UCF	8	0.310	22	Mississippi St	-1	0.047
(10-2)	23	UTSA	0.6968	+1.009		23	North Carolina	4	0.160	23	NC State	6	0.232
(8-4)	24	Ole Miss	0.6831	+0.938		24	Mississippi State	-3	0.134	24	Troy	-4	0.182
(10-2)	25	South Alabama	0.6792	+0.918		25	NC State	4	0.148	25	UTSA	-2	0.083
(9-2)	26	Coastal Carolina	0.6782	+0.913				calculated η →	1.167			calculated η →	1.275
(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			η (α)	η improvement			η (α)	η improvement	
(9-3)	31	Cincinnati	0.6404	+0.719		η (0)	1.167	-		η (0)	1.275	-	
(7-5)	32	Texas Tech	0.6361	+0.698									

Figure 7: 2022 week 13 - no MoV

Green cells are for user input

α	0.5

yellow cells are calculated

mean rating	0.500
std deviation	0.244

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	Committee/Poll Week 13		Diff to Aga-elo mat	| ln A(i)−ln C(i) |	Committee/Poll Week 14		Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(12-0)	1	Georgia	1.1751	+2.768		1	Georgia	0	0	1	Georgia	0	0
(12-0)	2	Michigan	1.0894	+2.416		2	Michigan	0	0	2	Michigan	0	0
(12-0)	3	TCU	1.0541	+2.272		3	TCU	0	0	3	TCU	0	0
(11-1)	4	Ohio State	1.0306	+2.175		4	USC	3	0.560	4	Ohio State	0	0
(10-2)	5	Tennessee	0.9779	+1.959		5	Ohio State	-1	0.223	5	Alabama	1	0.182
(10-2)	6	Alabama	0.9681	+1.919		6	Alabama	0	0	6	Tennessee	-1	0.182
(11-1)	7	USC	0.9674	+1.916		7	Tennessee	-2	0.336	7	Clemson	3	0.357
(10-2)	8	Penn State	0.9242	+1.739		8	Penn State	0	0	8	Utah	5	0.486
(9-3)	9	Kansas State	0.8693	+1.514		9	Clemson	1	0.105	9	Kansas State	0	0
(10-2)	10	Clemson	0.8566	+1.462	1.306	10	Kansas State	-1	0.105	10	USC	-3	0.357
(8-4)	11	Texas	0.8538	+1.451		11	Utah	2	0.167	11	Penn State	-3	0.318
(9-3)	12	LSU	0.8232	+1.325		12	Washington	4	0.288	12	Washington	4	0.288
(9-3)	13	Utah	0.8225	+1.322		13	Florida State	1	0.074	13	Florida State	1	0.074
(9-3)	14	Florida State	0.8219	+1.320		14	LSU	-2	0.154	14	Oregon State	4	0.251
(9-3)	15	Oregon	0.8173	+1.301		15	Oregon State	3	0.182	15	Oregon	0	0
(10-2)	16	Washington	0.8124	+1.281		16	Oregon	-1	0.065	16	Tulane	1	0.061
(10-2)	17	Tulane	0.7883	+1.182		17	UCLA	2	0.111	17	LSU	-5	0.348
(9-3)	18	Oregon State	0.7882	+1.181		18	Tulane	-1	0.057	18	UCLA	1	0.054
(9-3)	19	UCLA	0.7704	+1.109		19	South Carolina	8	0.351	19	South Carolin	8	0.351
(8-4)	20	Notre Dame	0.7504	+1.026		20	Texas	-9	0.598	20	Texas	-9	0.598
(8-4)	21	Mississippi State	0.7486	+1.019		21	Notre Dame	-1	0.049	21	Notre Dame	-1	0.049
(10-2)	22	Troy	0.7351	+0.964		22	UCF	4	0.167	22	Mississippi St	-1	0.047
(8-4)	23	Ole Miss	0.7274	+0.932		23	North Carolina	5	0.197	23	NC State	6	0.232
(10-2)	24	UTSA	0.7124	+0.871		24	Mississippi State	-3	0.134	24	Troy	-2	0.087
(10-2)	25	South Alabama	0.6963	+0.805		25	NC State	4	0.148	25	UTSA	-1	0.041
(9-3)	26	UCF	0.6951	+0.800				calculated η →	1.177			calculated η →	1.191
(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			η (α)	η improvement			η (α)	η improvement	
(9-3)	31	Cincinnati	0.6650	+0.677		η (0)	1.167	-		η (0)	1.275	-	
(7-5)	32	Louisville	0.6636	+0.671		η (.5)	1.177	-6%		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					

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 regression in η of -11%.

These results are typical -- η improvement of 20% to 50% can be seen in early and midseason weeks, and η improvement (regression) goes to -20 to -50% by the end of the season.

A way to read these results is that in the early season margin-of-victory heavily influences where teams deserve to be ranked, but in the late season who you beat matters more, at least for the purposes of the polls.

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[10], 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.

For example, let's look at the final ratings/rankings for 2022 and examine TCUs place. Here are the final results with both $\alpha = 0$ and $\alpha = 0.5,\ m \in (-0.25\ , 0\ ,\ 0.25)$

α	0		Aij adjustment	0
mean rating	0.500		Dii adjustment (not used)	0
std deviation	0.196		Committee / Poll Week	16

		team	rating			Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(15-0)	1	Georgia	1.0422	1	Georgia	0	0
(13-1)	2	Michigan	0.9031	2	TCU	1	0.405
(13-2)	3	TCU	0.8910	3	Michigan	-1	0.405
(11-2)	4	Ohio State	0.8665	4	Ohio State	0	0
(11-2)	5	Tennessee	0.8656	5	Alabama	1	0.182
(11-2)	6	Alabama	0.8579	6	Tennessee	-1	0.182
(11-2)	7	Penn State	0.8489	7	Penn State	0	0
(11-3)	8	USC	0.7985	8	Washington	4	0.405
(10-3)	9	Oregon	0.7906	9	Tulane	1	0.105
(12-2)	10	Tulane	0.7895	10	Utah	8	0.588

Figure 11: TCU with final ranking of #3 with no MoV

[10] 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.

				α	0.5		Aij adjustment	0.5	
		mean rating	0.500				Dii adjustment (not used)	0	
		std deviation	0.212						
						Committee / Poll Week	16		
		team	rating					Diff to Aga-elo matr	| ln A(i)−ln C(i) |
(15-0)	1	Georgia	1.1051			1	Georgia	0	0
(11-2)	2	Ohio State	0.9744			2	TCU	5	1.253
(13-1)	3	Michigan	0.9431			3	Michigan	0	0
(11-2)	4	Penn State	0.9136			4	Ohio State	-2	0.693
(11-2)	5	Alabama	0.9097			5	Alabama	0	0
(11-2)	6	Tennessee	0.8959			6	Tennessee	0	0
(13-2)	7	TCU	0.8876			7	Penn State	-3	0.560
(11-3)	8	USC	0.8191			8	Washington	10	0.811
(11-3)	9	Clemson	0.8057			9	Tulane	4	0.368
(10-4)	10	Kansas State	0.8038			10	Utah	4	0.336

Figure 12: TCU with ranking of #7 with MoV taken into consideration

We see TCU drops all the way to #7 in the playoffPredictor computer model. Is this correct? According to the pollsters, no. The AP put TCU #2 in the final poll, despite a 58-point drubbing to Georgia in the national title game. My estimation is that #2 spot is not correct. If they were going to play a follow on game against any of the teams ranked 1-6 they would be a significant Vegas underdog. The public just understands that TCU may deserve the #2 ranking in the final poll, but TCU is not the 2nd best team in the country.

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:

		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 would.

Note that this is not, and can not be a perfect model for every situation. For example, if a 12-0 BYU team beats all their WAC opponents by 1-point, yet still gets matched in the bowl game against Michigan and defeats Michigan by a small margin, there is a solid argument for undefeated and best in the country, but their strength-of-schedule and margin-of-victory would be so small that it is likely other teams with a loss would be rated ahead.

Another example would be if at the end of a season the #1 and #2 teams played each other a ridiculous number of times, like 100 rematches. Each team would end up with a record around .500, even though they were clearly the 2 best at the end of the regular season.

This method is not intended to be a perfect model for any hypothetical situation, but it is the best model I know of for the task. In fact, according to the Arrow impossibility therom there is no perfect model, and he won a Nobel prize for that insight.

SECTION 2

USING BIAS TO PREDICT COMMITTEE RANKINGS

7.0 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 computer ranked #1 and 2015 Michigan State at week 14, with a 12-1 record and also computer 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

8.0 Using ratings to compute future probabilities

8.1 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 $\sigma \approx 0.25$

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 at the end of the 2021 season were 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	***50.01%***
0.5	50.50%
1	51.20%
1.5	52.50%
2	53.40%
2.5	54.20%
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	75.00%
9.5	75.50%
10	77.40%
10.5	79.20%
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	87.40%
15	88.50%
16	89.00%
17	91.40%
18	93.70%
19	94.00%
20	94.70%
21	94.90%
22	95.50%
23	96.20%
24	96.40%
25	96.60%
26	96.80%
27	97.10%
28	97.30%
29	97.50%
30	97.80%
31	98.00%
32	98.20%
33	98.40%
34	98.60%
35	98.80%
36	99.00%
37	99.40%
38	99.70%
39	99.90%
40	100.00%

The range is from a -40 point home favorite to a +40 point road favorite. It is intentional to have more probability along multiples of 7 and 7x+3.

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) = 1.99% selects -32 point spread, just like P(Road Team Wins) = 2.00% also selects -32, whereas P(Road Team Wins) = 2.01% selects -31 point spread.

To test the validity of this method, we can look at a sample from 2022 predicting week 14 – conference championship week. That week has only 11 games contested, so it is fairly easy to see how the model works for a specific week. First off, we need to see what the model rates each team inputting only the first 13 weeks of data. We get this top 25 and which has an η=1.247:

					Committee/Poll Week		13		
		team	rating	Z*-score				Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(12-0)	1	Georgia	1.0412	+2.571		1	Georgia	0	0
(11-1)	2	Ohio State	0.9621	+2.196		2	Michigan	2	0.693
(12-0)	3	TCU	0.9506	+2.141		3	TCU	0	0
(12-0)	4	Michigan	0.9505	+2.140		4	USC	1	0.223
(11-1)	5	USC	0.8821	+1.815		5	Ohio State	-3	0.916
(10-2)	6	Penn State	0.8784	+1.798		6	Alabama	2	0.288
(10-2)	7	Tennessee	0.8614	+1.717		7	Tennessee	0	0
(10-2)	8	Alabama	0.8609	+1.715		8	Penn State	-2	0.288
(9-3)	9	Kansas State	0.8432	+1.631		9	Clemson	2	0.201
(8-4)	10	Texas	0.8266	+1.552		10	Kansas State	-1	0.105
(10-2)	11	Clemson	0.7993	+1.422		11	Utah	1	0.087
(9-3)	12	Utah	0.7870	+1.364		12	Washington	5	0.348
(9-3)	13	Florida State	0.7839	+1.349		13	Florida State	0	0
(9-3)	14	Oregon	0.7717	+1.291		14	LSU	1	0.069
(9-3)	15	LSU	0.7659	+1.263		15	Oregon State	4	0.236
(9-3)	16	UCLA	0.7496	+1.186		16	Oregon	-2	0.134
(10-2)	17	Washington	0.7469	+1.173		17	UCLA	-1	0.061
(10-2)	18	Tulane	0.7467	+1.172		18	Tulane	0	0
(9-3)	19	Oregon State	0.7405	+1.143		19	South Carolina	6	0.274
(8-4)	20	Notre Dame	0.7071	+0.984		20	Texas	-10	0.693
(8-4)	21	Mississippi State	0.6957	+0.930		21	Notre Dame	-1	0.049
(8-4)	22	Ole Miss	0.6871	+0.889		22	UCF	5	0.205
(10-2)	23	Troy	0.6870	+0.888		23	North Carolina	7	0.266
(10-2)	24	UTSA	0.6799	+0.855		24	Mississippi State	-3	0.134
(8-4)	25	South Carolina	0.6630	+0.774		25	NC State	7	0.247
(10-2)	26	South Alabama	0.6619	+0.769				calculated η →	1.247

Using those inputs for team rating we get the following predictions for week 14 games. The below table lists the 11 games played on the weekend of December 3, 2022, with the computer predicted winner and line, the Vegas line, and the actual score differential.

	base used->	1000							Computer			Vegas			
	home adv->	0				my predicted point spread	vegas point spread	Actual	Straight up win percentage (prediction)	Spread win percentage (prediction)	average of (Absolute value of (prediction - actual))	Straight up win percentage (Vegas)	Spread win percentage (Vegas) [0-0-N by definition]	average of (Absolute value of (prediction - actual))	Improvement over Vegas
	Multiplier ->	1				-7.32	-5.14		0.64	0.45	13.41	0.82	0.50	11.86	-1.55
									(7 - 4)	(4 - 5 - 2)		(9 - 2)	(0 - 0 - 11)		
home Conference	home	away	my rating home	my rating away	percentage away			Over Expected ->	*-0.182*	*-0.045*	*1.55*				
1-SEC	**Georgia**	**LSU**	1.04	0.77	12.99%	-14.5	-17	-20	1	0	5.5	1	0.5	3	-2.5
2-Big10	**Michigan**	**Purdue**	0.95	0.61	8.83%	-17	-15	-21	1	1	4	1	0.5	6	+2.0
3-B12	**TCU**	**Kansas State**	0.95	0.84	32.25%	-7	1	3	0	0	10	1	0.5	2	-8.0
4-ACC	**North Carolina**	**Clemson**	0.65	0.80	73.61%	8	7	29	1	1	21	1	0.5	22	+1.0
5-Pac12	**USC**	**Utah**	0.88	0.79	34.15%	-6	-3	23	0	0	29	0	0.5	26	-3.0
G5	**Buffalo**	**Akron**	0.36	0.15	18.77%	-12	-12	-1	1	0.5	11	1	0.5	11	0
G5	**UTSA**	**North Texas**	0.68	0.47	18.95%	-12	-8	-21	1	1	9	1	0.5	13	+4.0
G5	**Ohio**	**Toledo**	0.55	0.44	32.27%	-7	3.5	10	0	0	17	1	0.5	6.5	-10.5
G5	**Troy**	**Coastal Carolina**	0.69	0.61	37.24%	-4.5	-6.5	-19	1	0	14.5	1	0.5	12.5	-2.0
G5	**Boise State**	**Fresno State**	0.60	0.56	43.08%	-3	-3	12	0	0.5	15	0	0.5	15	0
G5	**Tulane**	**UCF**	0.75	0.66	35.41%	-5.5	-3.5	-17	1	1	11.5	1	0.5	13.5	+2.0

As we can see the model lines up sensibly with Vegas data. Only 2 of the 11 games are a greater discrepancy than 4 points, and 2 games (Buffalo-Akron and Boise State-Fresno State) line up exactly with the Vegas line.

The model has an average predicted point spread of -7.32, close to Vegas' -5.14. Furthermore, with regard to actual results, this sample has a spread win percentage of 45%, and mean average error (MAE) of 13.41. Vegas had a MAE of 11.86 for these games.

So we have sane data, but can we make it better? Let's look at how the data changes if we use an $\alpha = 0$,[11] which makes the input ratings Colley ratings:

[11] or $m = 0$, either way $\alpha m = 0$

					Committee/Poll Week	13			
		team	rating	Z*-score				Diff to Aga-elo mat	$\lvert \ln A(i) - \ln C(i) \rvert$
(12-0)	1	Georgia	0.9762	+2.440		1	Georgia	0	0
(12-0)	2	TCU	0.9456	+2.284		2	Michigan	1	0.405
(12-0)	3	Michigan	0.9196	+2.150		3	TCU	-1	0.405
(11-1)	4	USC	0.8789	+1.942		4	USC	0	0
(11-1)	5	Ohio State	0.8692	+1.892		5	Ohio State	0	0
(10-2)	6	Tennessee	0.8319	+1.701		6	Alabama	1	0.154
(10-2)	7	Alabama	0.8267	+1.674		7	Tennessee	-1	0.154
(10-2)	8	Penn State	0.8149	+1.614		8	Penn State	0	0
(10-2)	9	Clemson	0.7941	+1.507		9	Clemson	0	0
(9-3)	10	Kansas State	0.7740	+1.404		10	Kansas State	0	0
(9-3)	11	Oregon	0.7635	+1.350		11	Utah	6	0.435
(10-2)	12	Washington	0.7610	+1.337		12	Washington	0	0
(9-3)	13	LSU	0.7532	+1.298		13	Florida State	3	0.208
(8-4)	14	Texas	0.7428	+1.244		14	LSU	-1	0.074
(9-3)	15	Oregon State	0.7415	+1.238		15	Oregon State	0	0
(9-3)	16	Florida State	0.7394	+1.227		16	Oregon	-5	0.375
(9-3)	17	Utah	0.7351	+1.205		17	UCLA	2	0.111
(10-2)	18	Tulane	0.7307	+1.182		18	Tulane	0	0
(9-3)	19	UCLA	0.7280	+1.169		19	South Carolina	9	0.388
(10-2)	20	Troy	0.7234	+1.145		20	Texas	-6	0.357
(8-4)	21	Mississippi State	0.7027	+1.039		21	Notre Dame	1	0.047
(8-4)	22	Notre Dame	0.6969	+1.009		22	UCF	8	0.310
(10-2)	23	UTSA	0.6968	+1.009		23	North Carolina	4	0.160
(8-4)	24	Ole Miss	0.6831	+0.938		24	Mississippi State	-3	0.134
(10-2)	25	South Alabama	0.6792	+0.918		25	NC State	4	0.148
(9-2)	26	Coastal Carolina	0.6782	+0.913				calculated η →	1.167
(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			η (α)	η improvement	
(9-3)	31	Cincinnati	0.6404	+0.719		η (0)	1.167	-	
(7-5)	32	Texas Tech	0.6361	+0.698		η (.5)	1.247	-48%	

Again, we have used inputs of scores from weeks 1-13 in 2022. Now η has improved to η=1.167, which is a vast improvement of almost 50% (η=1 is perfect correlation). But how do these $\alpha = 0$ ratings compare to predictions for week 14?

									Computer			Vegas			
	base used->	1000													
	home adv->	0				my predicted point spread	vegas point spread	Actual	Straight up win percentage (prediction)	Spread win percentage (prediction)	average of (Absolute value of (prediction - actual))	Straight up win percentage (Vegas)	Spread win percentage (Vegas) [0-0-N by definition]	average of (Absolute value of (prediction - actual))	Improvement over Vegas
	Multiplier ->	1				-7.68	-5.14		0.64	0.36	14.23	0.82	0.50	11.86	-2.36
									(7 - 4)	(3 - 6 - 2)		(9 - 2)	(0 - 0 - 11)		
home Conference	home	away	my rating home	my rating away	percentage away			Over Expected ->	-0.182	-0.136	2.36				
1-SEC	Georgia	LSU	0.98	0.75	17.65%	-12.5	-17	-20	1	0	7.5	1	0.5	3	-4.5
2-Big10	Michigan	Purdue	0.92	0.61	10.59%	-17	-15	-21	1	1	4	1	0.5	6	+2.0
3-B12	TCU	Kansas State	0.95	0.77	23.40%	-10	1	3	0	0	13	1	0.5	2	-11.0
4-ACC	North Carolina	Clemson	0.68	0.79	69.16%	7	7	29	1	0.5	22	1	0.5	22	0
5-Pac12	USC	Utah	0.88	0.74	27.03%	-7.5	-3	23	0	0	30.5	0	0.5	26	-4.5
G5	Buffalo	Akron	0.37	0.16	18.90%	-12	-12	-1	1	0.5	11	1	0.5	11	0
G5	UTSA	North Texas	0.70	0.47	17.17%	-13	-8	-21	1	1	8	1	0.5	13	+5.0
G5	Ohio	Toledo	0.59	0.45	27.08%	-7.5	3.5	10	0	0	17.5	1	0.5	6.5	-11.0
G5	Troy	Coastal Carolina	0.72	0.68	42.26%	-3.5	-6.5	-19	1	0	15.5	1	0.5	12.5	-3.0
G5	Boise State	Fresno State	0.62	0.57	41.20%	-3.5	-3	12	0	0	15.5	0	0.5	15	-0.5
G5	Tulane	UCF	0.73	0.65	36.88%	-5	-3.5	-17	1	1	12	1	0.5	13.5	+1.5

Not so good. We have gone from a win against-the-spread rate of 45% (4-5-2) to 36% (3-6-2). The MAE (Mean Average Error) has increased by almost a full point from 13.41 to 14.23. So the net result is even though the poll is much more correctly predicted by $\alpha = 0$ weights, the actual game results are more correctly predicted by $\alpha = 0.5$ weights. What we see here is a divergence between most deserving ($\alpha = 0$) and best ($\alpha = 0.5$) teams. The Colley Matrix selects the most deserving. The PlayoffPredictor.com method selects a balance between the two.

To illustrate further, let's use the following maximum theoretical weights for games played through week 13:

$$m = \begin{cases} -1, & for\ MoV\ of\ 1-7\ points \\ 0, & for\ MoV\ of\ 8-24\ points \\ +1, & for\ MoV\ of\ 25+\ points \end{cases}$$

Remember, -1 is like a win by default (both teams A&B wind up with 0.5 ratings) and +1 is like A beats B every time in infinite rematches.

Using that input we get the following rating / ranking:

					Committee/Poll Week	13			
		team	rating	Z*-score				Diff to Aga-elo mat	| ln A(i)−ln C(i) |
(11-1)	1	Ohio State	1.3271	+2.995		1	Georgia	1	0.693
(12-0)	2	Georgia	1.2555	+2.736		2	Michigan	4	1.099
(8-4)	3	Texas	1.1262	+2.268		3	TCU	4	0.847
(10-2)	4	Penn State	1.1023	+2.181		4	USC	8	1.099
(9-3)	5	Kansas State	1.0924	+2.146		5	Ohio State	-4	1.609
(12-0)	6	Michigan	1.0428	+1.966		6	Alabama	3	0.405
(12-0)	7	TCU	0.9729	+1.713		7	Tennessee	3	0.357
(9-3)	8	Utah	0.9623	+1.674		8	Penn State	-4	0.693
(10-2)	9	Alabama	0.9484	+1.624		9	Clemson	9	0.693
(10-2)	10	Tennessee	0.9398	+1.593		10	Kansas State	-5	0.693
(9-3)	11	Florida State	0.9071	+1.474		11	Utah	-3	0.318
(11-1)	12	USC	0.8859	+1.398		12	Washington	18	0.916
(6-6)	13	Oklahoma	0.8454	+1.251		13	Florida State	-2	0.167
(8-4)	14	Illinois	0.8151	+1.141		14	LSU	2	0.134
(9-3)	15	UCLA	0.8067	+1.111		15	Oregon State	7	0.383
(9-3)	16	LSU	0.7952	+1.069		16	Oregon	4	0.223
(4-8)	17	Iowa State	0.7930	+1.061		17	UCLA	-2	0.125
(10-2)	18	Clemson	0.7878	+1.042		18	Tulane	1	0.054
(10-2)	19	Tulane	0.7818	+1.021		19	South Carolina	29	0.927
(9-3)	20	Oregon	0.7687	+0.973		20	Texas	-17	1.897
(7-5)	21	Louisville	0.7645	+0.958		21	Notre Dame	3	0.134
(9-3)	22	Oregon State	0.7248	+0.814		22	UCF	11	0.405
(7-5)	23	Wake Forest	0.7240	+0.811		23	North Carolina	35	0.925
(8-4)	24	Notre Dame	0.7238	+0.811		24	Mississippi State	10	0.348
(6-6)	25	Arkansas	0.7215	+0.802		25	NC State	35	0.875
(6-6)	26	Florida	0.7192	+0.794				calculated η →	1.898

η has decreased all the way to η=1.898, which is awful, especially late in the season. Notice Texas is the #3 team, even though they have 4 losses. But all those losses were less than 7 points, so now $m = -1$ losses. Likewise, 12-0 Michigan is at #6 way behind 11-1 Ohio State at #1, who was just beaten by Michigan in week 13. Why? Because in those other 22 games — Ohio State blew out 5 teams ranked 44-91. Michigan blew out 5 teams also, but those teams were ranked 90-120. The math, when emphasizing propensity to blowout competition, works out to Ohio State's strong favor, even with the solid 22-point loss to Michigan.

But again, the real question is, how do these αm ratings manifest as predictions for week 14?

	base used->	1000							Computer			Vegas			
	home adv->	0				my predicted point spread	vegas point spread	Actual	Straight up win percentage (prediction)	Spread win percentage (prediction)	average of (Absolute value of (prediction - actual))	Straight up win percentage (Vegas)	Spread win percentage (Vegas) [0-0-N by definition]	average of (Absolute value of (prediction - actual))	Improvement over Vegas
	Multiplier ->	1				-5.55	-5.14		0.82	0.82	10.45	0.82	0.50	11.86	1.41
									(9 - 2)	(8 - 1 - 2)		(9 - 2)	(0 - 0 - 11)		
home Conference	home	away	my rating home	my rating away	percentage away			Over Expected ->	*0.000*	*0.318*	*-1.41*				
1-SEC	**Georgia**	**LSU**	1.26	0.80	3.99%	-23	-17	-20	1	1	3	1	0.5	3	0
2-Big10	**Michigan**	**Purdue**	1.04	0.62	4.98%	-22	-15	-21	1	1	1	1	0.5	6	+5.0
3-B12	**TCU**	**Kansas State**	0.97	1.09	69.54%	7	1	3	1	1	4	1	0.5	2	-2.0
4-ACC	**North Carolina**	**Clemson**	0.56	0.79	83.11%	13.5	7	29	1	1	15.5	1	0.5	22	+6.5
5-Pac12	**USC**	**Utah**	0.89	0.96	62.88%	4.5	-3	23	1	1	18.5	0	0.5	26	+7.5
G5	**Buffalo**	**Akron**	0.31	0.10	18.42%	-12	-12	-1	1	0.5	11	1	0.5	11	0
G5	**UTSA**	**North Texas**	0.62	0.46	25.09%	-9	-8	-21	1	1	12	1	0.5	13	+1.0
G5	**Ohio**	**Toledo**	0.43	0.41	44.93%	-3	3.5	10	0	0	13	1	0.5	6.5	-6.5
G5	**Troy**	**Coastal Carolina**	0.58	0.45	28.73%	-7.5	-6.5	-19	1	1	11.5	1	0.5	12.5	+1.0
G5	**Boise State**	**Fresno State**	0.55	0.52	45.36%	-3	-3	12	0	0.5	15	0	0.5	15	0
G5	**Tulane**	**UCF**	0.78	0.68	32.79%	-6.5	-3.5	-17	1	1	10.5	1	0.5	13.5	+3.0

As you can see, fantastic! We have gone from a win against-the-spread rate of 45% (4-5-2) to 82% (8-1-2). The MAE (Mean Average Error) has decreased all the way to 10.45, almost 1.5 under Vegas. Again, we see divergence between most deserving ($\alpha = 0$) and best ($\alpha = 0.5, m = +1, -1$) teams. The selecting of weights of $m = +0.25, -.25$ is a good balance between most deserving and best, with the best results to historical backtesting.

Of course, the model does not get ATS results around 80% for large sample sizes, even if you use the weights of $m \in (+1, 0, -1)$. Using those weights for the whole season of 2022 you get an ATS of 49.3%. However, the basic point that as m increases from $m = 0$ to $m \in (+0.25, 0, -0.25)$ the general trend is for the η to get worse and the MAE & ATS to get better (closer to Vegas or potentially above Vegas)

And here we come to a critical point, perhaps the most important result to fall out of this whole method. Namely:

Axiom #3: There are distinctly different equations for generating m margin for arriving at the most deserving teams and the best teams

The best teams $\neq$ the most deserving teams. This is not just a hunch; it is backed by data. When we use an m generating function that does not punish close wins (or reward close losses) we get an η that comes very close to the AP and committee polls, but does not do as well against the spread and against a mean average error statistic compared to the Vegas line. When we use an m generating function that does seriously diminish the value of close wins and gives extra rewards for lopsided blowouts we get a result that can be used to beat the Vegas spread and produces a mean average error much closer to the Vegas line, but increases η very significantly and begins to diverge with the AP and committee polls.

.

8. 2 Results for full season with PlayoffPredictor ATS weights

We now backtest the method and see results for the 2022 season. Lines were taken from api.collegefootballdata.com and are closing lines (just before kickoff) from the 1st returned provider of lines.

2022 playoffPredictor weekly η poll results and ATS betting results

	Vegas	PlayoffPredictor model with m=0.25		# of games	won games (m=0.25)
week 0 η					
week 1 ATS / 1000,0.05,1	*50%*	43%		94	40.5
week 1 MAE / 1000,0.05,1	11.51	24.29			
week 1 η		2.73			
week 2 ATS / 1000,0.05,1	*50%*	54%		80	43.0
week 2 MAE / 1000,0.05,1	13.04	15.53			
week 2 η		2.78			
week 3 ATS / 1000,0.05,1	*50%*	51%		72	36.5
week 3 MAE / 1000,0.05,1	13.01	18.42			
week 3 η		2.44			
week 4 ATS / 1000,0.05,1	*50%*	48%		66	32.0
week 4 MAE / 1000,0.05,1	12.05	14.48			
week 4 η		2.00			
week 5 ATS / 1000,0.05,1	*50%*	53%		63	33.5
week 5 MAE / 1000,0.05,1	11.40	12.38			
week 5 η		1.716			
week 6 ATS / 1000,0.05,1	*50%*	45%		57	25.5
week 6 MAE / 1000,0.05,1	11.98	13.14			
week 6 η		1.419			
week 7 ATS / 1000,0.05,1	*50%*	54%		53	28.5
week 7 MAE / 1000,0.05,1	12.3	13.93			
week 7 η		1.321			
week 8 ATS / 1000,0.05,1	*50%*	52%		53	27.5
week 8 MAE / 1000,0.05,1	11.31	11.87			
week 8 η		1.389			
week 9 ATS / 1000,0.05,1	*50%*	51%		48	24.5
week 9 MAE / 1000,0.05,1	12.59	13.26			
week 9 η		1.356			
week 10 ATS / 1000,0.05,1	*50%*	55%		60	33.0
week 10 MAE / 1000,0.05,	11.71	12.54			
week 10 η		1.295			
week 11 ATS / 1000,0.05,1	*50%*	47%		57	27.0
week 11 MAE / 1000,0.05,	12.9	13.85			
week 11 η		1.265			
week 12 ATS / 1000,0.05,1	*50%*	52%		59	30.5
week 12 MAE / 1000,0.05,	12.92	13.22			
week 12 η		1.22			
week 13 ATS / 1000,0.05,1	*50%*	43%		63	27.0
week 13 MAE / 1000,0.05,	12.58	13.29			
week 13 η		1.25			
week 14 ATS / 1000,0.05,1	*50%*	45%		12	5.5
week 14 MAE / 1000,0.05,	11.86	13.41			
week 14 η		1.298			
bowls ATS / 1000,0.05,1	*50%*	44%		42	18.5
bowls MAE / 1000,0.05,1	12.19	12.82			
end-of-season η		1.33			

total games played and won ATS - full season	879	433.0
total games played and won ATS - weeks 2-16	785	392.5

season final ATS %	49.26%
weeks 2-16 ATS %	50.00%

We can observe some definite results here. In week 0 (predicting week 1) we have an awful MAE of 24.29. Of course it is going to be awful – all teams are tied for 1st place with a 0.500 rating. Every single game is going to be a predicted point spread of -3.5 (because of the home team advantage of 0.05). Predicting week 1 with this model is not really a prediction that makes any sense. The point of this model is that results on the field map to ratings and rankings. Just saying every home team is favored by 3.5 points is as likely to result in success as saying every home team will cover or every favorite will lose against the spread. They are strategies, but not so much the output of a system of equations with tested weights. It is more meaningful to look at results from weeks 2 through the end of the season.

Weeks 2-11 are all relatively good, several weeks of 55% ATS, with only 3 weeks under 50%.

Weeks 12-14 (predicting rivalry week, conference championship week, and bowls) are not good – an ATS of 44%. But remember the results for those weeks can be excellent if the weights are changed closer to $m \in (+1, 0, -1)$.

Every single week has a MAE worse than Vegas.

Overall, with weights of $m \in (+0.25, 0, -.25)$ for the range of $MoV \in (1-7, 8-24, 25+)$, home field advantage $h = 0.05$ and b base of 1000 we get a week 2-16 cumulative ATS of exactly 50% picking a total of 785 games.

High-confidence picks

That is the result of the computer picking all 879 games in 2022. How about identifying where the model sees something Vegas does not? If we only pick games where the Vegas line and the playoffPredictor.com line disagree by more than 7 points we see in the next chart how full season ATS numbers come out:

2022 playoffPredictor weekly ATS results	Vegas	m=0.25 / high confidence picks	# high conf games	won games hcp
week 0 η				
week 1 ATS / 1000,0.05,1	*50%*	36%	60	21.5
week 1 MAE / 1000,0.05,1	11.51			
week 1 η				
week 2 ATS / 1000,0.05,1	*50%*	58%	53	30.5
week 2 MAE / 1000,0.05,1	13.04			
week 2 η				
week 3 ATS / 1000,0.05,1	*50%*	46%	48	22
week 3 MAE / 1000,0.05,1	13.01			
week 3 η				
week 4 ATS / 1000,0.05,1	*50%*	47%	33	15.5
week 4 MAE / 1000,0.05,1	12.05			
week 4 η				
week 5 ATS / 1000,0.05,1	*50%*	60%	21	12.5
week 5 MAE / 1000,0.05,1	11.40			
week 5 η				
week 6 ATS / 1000,0.05,1	*50%*	50%	17	8.5
week 6 MAE / 1000,0.05,1	11.98			
week 6 η				
week 7 ATS / 1000,0.05,1	*50%*	63%	16	10
week 7 MAE / 1000,0.05,1	12.3			
week 7 η				
week 8 ATS / 1000,0.05,1	*50%*	63%	19	12
week 8 MAE / 1000,0.05,1	11.31			
week 8 η				
week 9 ATS / 1000,0.05,1	*50%*	65%	13	8.5
week 9 MAE / 1000,0.05,1	12.59			
week 9 η				
week 10 ATS / 1000,0.05,1	*50%*	62%	13	8
week 10 MAE / 1000,0.05,	11.71			
week 10 η				
week 11 ATS / 1000,0.05,1	*50%*	33%	12	4
week 11 MAE / 1000,0.05,	12.9			
week 11 η				
week 12 ATS / 1000,0.05,1	*50%*	36%	11	4
week 12 MAE / 1000,0.05,	12.92			
week 12 η				
week 13 ATS / 1000,0.05,1	*50%*	40%	10	4
week 13 MAE / 1000,0.05,	12.58			
week 13 η				
week 14 ATS / 1000,0.05,1	*50%*	0%	2	0
week 14 MAE / 1000,0.05,	11.86			
week 14 η				
bowls ATS / 1000,0.05,1	*50%*	25%	4	1
bowls MAE / 1000,0.05,1	12.19			
end-of-season η				
total games played and won ATS - full season			332	162.0
total games played and won ATS - weeks 2-16			272	140.5
	season final ATS %			48.80%
	weeks 2-16 ATS %			51.65%

We see much better results, especially in predicting weeks 2-10. We get weekly win rates against-the-spread of up to 65%. Things get much worse in predicting weeks 11 through the bowl season, dropping to a low of 0% (0 correct out of 2 total picks) for week 14.

Overall, with weights of $m \in (+0.25, 0, -.25)$ for the range of $MoV \in (1-7, 8-24, 25+)$, home field advantage $h = 0.05$ and b base of 1000 and only picking games where the Vegas line differs from the playoffPredictor.com line greater than 7 points, we get a week 2-16 cumulative ATS of 51.65% picking a total of 272 games.

Is that good? More importantly, is it just simple luck? I would say it is not good enough. A good system should have an ATS>55% with n>200 games picked. As we will see in the next section ATS numbers >65% are actually impossible.

8.3 Results with perfect forward information

Instead of backtesting, let's consider perfect forward information. If we had the final season ratings for every team, how good would the model do against the spread?

Computer					Vegas		
Straight up win percentage (prediction)	**Spread win percentage (prediction)**	**MAE: average of (Absolute value of (prediction - actual))**	**High confidence pick ATS : abs(Vegas-pp) >= 7.5**		**Straight up win percentage (Vegas)**	**Spread win percentage (Vegas) [0-0-N by definition]**	**MAE: average of (Absolute value of (prediction - actual))**
0.81	**0.65**	**11.26**	**0.74**		**0.74**	**0.50**	**12.21**
(705 - 163)	(548 - 284 - 36)		297		(638 - 230)	(0 - 0 - 868)	
0.077	***0.152***	***-0.96***					

What we see is that with even with perfect information we can only beat the spread 65% of the time. Even high confidence picks are only right 74% of the time. An example of a high confidence pick with perfect forward information would be picking TCU-Oklahoma in week 5. The final rankings had TCU at #7 and Oklahoma at #46, but since this was not known at the time of the game Vegas had TCU a +5 point underdog at home. The playoffPredictor model computes that game as -18 point favorite for TCU when given final season ratings. The actual final margin was TCU -31. It looks like a no-brainer in the rear-view mirror, but high confidence picks are only correct less than three-quarters of the time!

Conclusions

We have come up with a system that is simple, open, no prior-season information that produces great results vetted against real backtesting.

There is a difference between most deserving and best. Most deserving correlates to low η and is produced with weights for m that do not penalize close win (wins by 3-10 points). Best correlates to performance ATS with low MAE and is produced with weights for m that do penalize close wins and reward blowouts.

It is my hope people will use the principles, methods, and metrics of this paper and expand their own systems to get ATS >55% and MAE under the Vegas line.