

Simulations for the EPL Using Competitive Balance Models

G. Donald Allen

Department of Mathematics Texas A&M University College Station, TX 77843

Abstract: *In this paper we consider probabilistic competitive balance (CB) models for how football teams perform in English Premier League (EPL) play. While individual game outcomes are measured, they serve to only create the league table, and then over multiple seasons. This is a comprehensive model that attempts to model outcomes to accurately reflect actual league performance. For example, how does the top ranked team perform during a given season? Or over many seasons? Can the lower ranked teams do well or even win the championship? What is the shape of the points distribution over a simulated season? To make this model, we consider three fundamental factors, team skill, number of significant events during a match, and a criteria to win an event, lose, or tie a match. We then simulate the play focusing on the significant events relative to any two contestants.*

Keywords: *Simulation, Football modeling, Prediction*

I. Introduction

With only a few parameters in the model, which require considerable tuning, how can it be that the myriad of factors affecting team outcomes can be accurately modeled? We basically only utilize team skill levels and for an individual game the number of significant events that occur for pairs of teams. With just these factors, the model produces results remarkably similar to actual outcomes. Particularly interesting are the mid range teams that show a remarkable variation in their league outcomes as compared with both top and bottom ranked teams. We also compare the spread of winning percentage from the bottom to the top team and show the model accurately predicts it to be in the 40-60% range. Naturally, when the league champion and top ranked teams have rather high winning percentages the lowest and relegated teams have a rather low winning percentages, or points. Factors for professional sports leagues are remarkably similar, though league play varies considerably. There are common desiderata. The idea is that outcomes should be mildly but not strictly predictable. At the beginning of a season, we often read, "All the experts predict Team A should win the championship. Team B has scarcely a chance to do well." Only a minority are disappointed when either don't. During the course of each match there is a definite but highly complex random behavior, not nearly as simply as flipping a coin. There are dozens of nuances that affect games, and a model applying even few dozen factors probably will not achieve results comparable with experts such as bookmakers [1]. A generalized probabilistic model can usually do as well, though not match-to-match. It seems a natural human trait not to expect the top team to win and the bottom team to do better than expected. We separate amateur (e.g. college) and professional teams, as the goals are quite different. The fans of the fictitious Dickensonian Small College that has a final win-tie-loss record of 2-2-21 diminishes the value only for the season, but not so much the enthusiasm of the fans, or compel a coaching change. Indeed, we do see in amateur college sports extremely wide variations between winning and losing percentages within and across divisions. Here are some other expectations, at least in professional sports. Any model must accommodate these types of outcomes.

1. The season outcome should not be determined. There should be wide latitude from the league winner and that of other teams. Even the most favored team should have only about 40-60% chance of winning the league championship.
2. The difference between winners and losers should be significant but not ridiculously high. About a 35-55% spread in percentage between league winner and loser is acceptable. For a sport with multiple outcomes as ties, the notion of percentage wins needs some clarification. For football the percentage is easily computed as $(3 * \text{wins} + \text{ties}) / \text{pts available}$. In other leagues, such as the Spanish *La Liga*, with three teams perennially dominant, the differential is higher. However, this spread holds with other sports leagues such as American baseball. In basketball, the spread varies widely from about 20-60%. In NCAA basketball the spread may be up to 80% and rarely below 50%. The latter reflects the nature of a limited talent pool, including coaching and the college's internal emphasis on winning. Models for various leagues depend on these factors, particularly a vector we call the skills profile.
3. The outcome of any particular match is undetermined, particularly when the teams are equally matched. Nonetheless, any model should illustrate the outcomes can go either way, even between the best and worst teams. This does indeed happen, and because it is so striking, we tend to remember it, forgetting how rare it may be. For example, in 2012 the semi-pro team (Luton) defeated EPL team Norwich.

The general outline of this paper follows traditional lines for simulations. After a discussion of relative team skills a new notion of significant events per match (Section 3) we define the competition matrix, creating the a set of probabilities of how one team interacts with another (Section 4) for a win, tie, or loss. For a league of 20 teams such as the EPL this would necessitate the assignment of 190 values. At this point we look at W-D-L data for the EPL over the recent seasons to assign skills to respective teams (Section 5). This assignment is rather a significant task. Then, we give a detailed description of the simulation model used here where another new notion of event horizon is introduced (Section 6). The core portion of the paper is the report out of results. In Section 7 we discuss the net outcomes from the model as statistically compared with actual EPL data. In addition, a matrix of frequencies of how the variously ranked teams perform and their final rank positions compared with their initial rank. Section 8, on conclusions, critiques the model and poses numerous questions toward its improvement.

II. A Brief Review of Literature

In professional sports, owners do all they can to support and improve their teams. They try to hire the best managers and players within their proclivities and budgets. These and other factors generate the competitive balance or imbalance between teams. In addition there are numerous other factors relevant to the league. The more competitive teams are, the more uncertain team outcomes become, and the greater the attendance and other audience may become. These are components of the “uncertainty of outcomes hypothesis,” [2], [3].. In short, predictability of outcomes is an anathema to financial success.

Quantities most measured include means, variation, and end-of-season outcomes [4], [5], [6]. Even the Gini coefficient of the wealth within a country, and also of point spread applied to point totals has been considered [7], [8], as well as other measures [9], [10], [11]. Yet each sports league has its own peculiarities that may not transfer easily from one league to another, much less from one sport to another. A careful study of individual leagues is important. Only with a highly accurate model of one league is it possible to understand how to adjust it to other. Moreover, quantities used should be measurable.

Many measures use notions of dispersion between different sports and the comparison via the RSA, or relative standard deviation [12]. Simulation models have been applied to different aspects of the analysis of sports leagues, often to make predictions [13], [14], including the effects of match attendance as a consequence changes in league structure such a applying a draft [15]. More recently, there are papers applying artificial intelligence primarily to predict game-day outcomes [16], [17]. These have met with some success, but lie beyond the scope of this work.

III. Team Skills and Game Events

Our population is all those playing football beyond the club level, semi-pro and pro. We assume that the skill level of each player is in the range (0,100), with the higher value indicating a greater skill. Assume their skills to be normally distributed with mean 50 and standard deviation (50/3), placing almost every player in the assigned range. We assume that most if not all professional players in the major football leagues have skills above at least one standard deviation above the mean, or have skill level greater than 66. That is, we assume the top professional teams get a rather high percentage of this group, and this percentage measures the overall team skill.

A cursory, though not detailed inspection, show that the better teams in the EPL have a higher percentage of international players, indicating they have recruited into the broader worldwide pool of quality professional players. This applies to the major European football leagues in countries with a higher Gini index, notably the EPL, but also the *Bundesliga*, *La Liga*, and others.

Yet, a model based on relative skill levels of the teams begs the questions.

1. How can correct skill levels for teams be derived? Though this may be theoretically difficult, it is possible to model. One possibility is for players to be given a point value, ala chess, together with the manager, and this used to create the skill fractionations.
2. In this context, what is the valuation for a manager in relative terms of that for the team? Arguably, the manager is likely the most important component of the enterprise. In fact, in the most recent 14-15 EPL season, one particular team jumped about 10% in points substantially by changing only the manager.

In this paper, we use the final point standings in the league to estimate relative skill levels in the range [65,90] as discussed just above. It is one thing to define the skills of particular teams by a single number, and this is what is done in most competitive balance models, but it is another to allow the skill level of a team to be given by a probability density. The latter could be more accurate and would allow more variation in competitive outcomes. Yet, even so, it becomes important to assign the means of these densities, and more complex to define models for the density, e.g. uniform, normal, or even beta. This is discussed further in Section 4.

Events. Each match is not so simple as a win or loss, or even win, loss, or tie. Additional factors are required to fold it into a realistically model an entire game, much less an entire season between teams. We posit that each game between any two reasonably matched teams such as in the EPL or even the EPL and the Champions league consists of relatively routine play. During this which is the majority of the game, both teams play about the same. In fact, a good semi-pro team could likely play even with the best EPL team even 90% of the time. This leads to the suggestion that match outcomes are determined by **significant events** during the match. The questions become then, what are these events? How many of them are there per game? How is the winner of an individual event determined? And how does the preponderance of events wins determine a win, tie, or loss of a game?

Table 3.1 Significant Events

Action	D	O
TOTAL	35	40
Effective clearing of the ball	10	
A good defensive save	7	
An important tackle near the goal	5	
A strategic foul	4	
Goalie stopping a certain score	3	
Defending a corner	3	
The corner kick	1	
Strategic substitution	2	2
A good pass for a break away		7
A strong give and go		7
An effective long pass		7
Dribbling in traffic		4
Defensive miscue		4
Blowing a certain shot on goal		4
A great shot on goal		3
Excellent centering of the ball on a corner		2

First, we create a table of events, as observed by the author over many years of playing football and observing the game. There is no claim to expertise, that being in the domain of the true experts. In Table 3.1 we give a guess of the base number of significant events per game. In this table “D” means a defensive play, and “O” means an offensive play. The plays and their counts are debatable, but it does suggest that in ninety minutes of continuous play there are about 75 truly significant event, and rarely do these involve a goal scored. However, in many if not most cases, a score results from at least one and likely from a sequence of significant events. This is well understood in American football and baseball.

To mention one of these categories only in deflection to objections, it is well known that the number of corner kicks leading directly to a goal is small, maybe about 3%, but the action of a goal kicks implies previous challenges to the goal, which in turn implies blown shots or good defensive plays. (See <http://www.soccerbythenumbers.com/2011/05/why-goal-value-of-corners-is-almost.html>)

Using a dash of common sense, we suggest that for teams closely matched there are more significant events in a match. The reasoning may seem counter intuitive. For CB (almost equal) teams, each will play their passes and saves within normal abilities. For mismatched teams, the significant plays for the better team will increase slightly, those for the lesser team will often be countered. The net result is there are fewer significant plays overall. In addition, between greatly mismatched teams the stronger team will not play with the same intensity against the weaker team, also with the net result of fewer significant events. In our calculations for the EPL the numbers drop at most about 10% during a game. Relative play is the nature of most contests, with expectation levels appending this.¹

Most scores are not dumb luck; therefore the upshot is for any good model, it is important to have a measure of significant plays. We have so far addressed the first two of the questions, with the others awaiting discussion. Not considered directly in this paper are long term events such as ball control, continuous offensive pressure, overall team physical condition, and constancy of playing intensity, all difficult qualities to quantify.

¹ In analogy with the game of chess, the grand master can usually defeat the lesser opponent while reading Kant’s *A Critique of Pure Reason*, but not so when playing an evenly matched opponent. Indeed, the master will not need as many significant moves, and the lesser player has fewer to make.

IV. The Competition Matrix

Let $S_i, i=1\dots n$ be the skill levels of the number of the league teams labeled $i=1\dots n$. In our present situation $n=20$. As noted skills are in the $(0,100)$ range. We define the **competition matrix** between teams of the various skill levels to be

$$c_{ij} = P(\text{Team } i \text{ defeats Team } j), 1 \leq i < j \leq n$$

where $P(\text{Team } i \text{ defeats Team } j)$ is the probability that Team i wins an event over Team j . To avoid probabilistic difficulties, we assume $P(\text{Team } i \text{ defeats Team } i) = 0.5$, and the **complementarity condition** $c_{ji} = 1 - P(\text{Team } i \text{ defeats Team } j) = c_{ij}, 1 \leq j < i \leq n$

Several possibilities for C are available. One is a model first suggested by Bradley and Terry (Bradley and Terry, 1952). It is essentially given by

$$c_{ij} = \frac{S_i}{S_i + S_j}, 1 \leq i < j \leq n$$

We will call this the basic **harmonic competition matrix**, or if you prefer the **Bradley-Terry matrix** [18]. It builds-in the complementarity condition. With a redefinition of the skills vector via a logarithmic transformation, this is equivalent to the model used by Bradley and Terry, namely $c_{ij} = \frac{e^{S_i}}{e^{S_i} + e^{S_j}}, 1 \leq i < j \leq n$.

These are the probabilities that Team i will succeed over Team j in a particular significant event during a game. Another model, and the one used in this study is the **geometric competition matrix** given by

$$c_{ij} = \frac{S_i}{2\sqrt{S_i S_j}}, 1 \leq i \leq j = 1, \dots, n$$

The complementarity condition gives the lower half of C . This matrix tends ever so slightly to exaggerate the difference between teams at the extremes of skill ranges. For nearly equal skills in our range the two matrices are imperceptibly close. This ever so slight change does make a difference in outcomes. Bear in mind, neither of these competition matrix models have a mathematical, statistical, theoretical corner upon the manner in which competition actually takes place, given the skills vector. They are, however, relevant exigencies for the model. In general, we could assume that each of the skill levels S_i is a random variable with of mean S_i pertaining to some probability density f_i . In this case [19]

$$c_{ij} = P(i | j) = \int_{-\infty}^{\infty} f_i(t) \int_{-\infty}^t f_j(s) ds dt, 1 \leq i, j \leq n$$

The complementarity condition is built-in. Note, as indicated above we may restrict the range of these densities to the range $(0,100)$. Here, it becomes possible to allow a fully general skill range over $(-\infty, \infty)$, with means in $(0,100)$. Assuming each S_i is normally distributed with specific standard deviations σ_i we can mimic the competition matrices we apply using the geometric formulation. However, this produces quite a different model, and requires a considerable discussion. What we will call the event horizons in [Section 6](#) is preserved but in another form. It is quite general and includes with the definition of suitable probability densities of the previous two examples. This model will be applied at another time.

Mathematically, competition matrices are examples of positive matrices, in the sense of having positive entries, which have rather interesting properties. To our knowledge this particular form for positive matrices has not been considered in the literature.

V. The Data

To set up our model, we now outline the data used. It is substantially based on the relative skill levels of league teams. The skill levels will be used to create a competition matrix. With this matrix, criteria for numbers of significant events per game, and criteria for ties, we model a season's play. The data set includes the nine 06-07 through 14-15 seasons of the EPL averages of wins, ties, and losses, which incidentally exhibit a remarkable consistency. One may posit that more seasons should be included. However, every league sees trends over time, and to inaccurately weight a long past season with those more current could lead to unnatural outcomes.

There is a need to fully consider the nature of ties, a particularly troublesome issue. We also need to understand how to model the team skill set, whether to use the averages from the most recent season, or to

modify it, the latter of team changes in the off-season. In this study, we only consider final ranks over all teams, not necessarily individual ranks. However, we do illustrate how the variously ranked teams fair with respect to final ranks over the 100 multiple seasons simulated. As it turns out the top ranked teams normally finish near the top, while the bottom ranked teams rarely rise even to the mid range of final standings.

What we do not consider here, but could be incorporated into the model is the more specific nature of home and away game results. However, we do accommodate the home and away differential in the criteria for win, lose, or tie for home and away games. Using actual home-away percentages directly would be a next step. However, the idea here is toward a form of minimalism. That is, we utilize the minimal data set to generate reasonable team differentials, which through the model generates excellent results. As noted previously, we are not performing game-day estimates about which teams win, lose, or tie, but the model could be used to that end. This could involve using far more data including an application of prospect theory which in this context would use results of recent matches [20], [21], [22]. See particularly the paper by Burgos. Note, these papers take some interpretation to put into competitive sports, but the basic idea is that a win of this week's game perhaps gives bump in confidence and by association team skills for the next game.

Table 4.1 EPL Cumulative Averages

Rank	Ave wins	Ave ties	Ave Points	Pct
1	26.89	6.33	87.00	0.76
2	24.78	7.89	82.22	0.72
3	22.56	8.11	75.78	0.66
4	20.56	10.00	71.67	0.63
5	18.56	9.89	65.56	0.58
6	17.22	10.56	62.22	0.55
7	16.00	10.33	58.33	0.51
8	14.67	10.00	54.00	0.47
9	13.44	10.78	51.11	0.45
10	13.44	8.33	48.67	0.43
11	11.56	11.78	46.44	0.41
12	11.56	10.33	45.00	0.39
13	10.89	10.33	43.00	0.38
14	10.44	10.11	41.44	0.36
15	10.44	9.00	40.33	0.35
16	8.78	11.89	38.22	0.34
17	9.00	10.11	37.11	0.33
18	8.67	9.22	35.22	0.31
19	7.67	9.67	32.67	0.29
20	5.78	9.56	26.89	0.24
Average	14.14	9.71	52.14	0.46
Hi	26.89	11.89	92.56	0.81
Lo	5.78	6.33	23.67	0.21
Standard Deviation	5.94	1.31	19.12	0.17
Ratio hi:lo	4.65	1.88	15.84	0.14
Ave top 3	24.74	7.44	81.67	0.72
Ave bottom 3	7.37	9.48	31.59	0.28

For football, we used to some advantage the percentage of points. This is the ratio of the number of points scored to the total points available, which in a 20-team league, each playing every other team twice gives a maximum number of points equal to thrice the number of games played (38), and this is 114. In Table 4.1 we show the average results over the nine seasons considered, 06-07 to 14-15. The spread is about 50 percentage points, normal in professional sports. From the top of the graph, we see the percentages of the top ranked teams jump somewhat more than the percentages near the bottom of the ranks. This also is generally the situation for individual seasons. This is something of a **point signature** for the league. For *La Liga*, where typically three teams dominate the game the percentages are somewhat higher, and in consequence the bottom-ranked teams have lower percentages. But the differentials for the top teams and bottom teams remain.

With this in mind, we graph the percentages or points by rank. Shown in Fig. 4.1, we have a striking visual of how teams perform by rank. Note, it is this shape, a slightly reverse sigmoidal curve that remains

consistent over the seasons. Now, we assign the skills values of the ranked teams in a linear scale from the number of points to the scale from the range [26.89, 87.00] to the range [65,90]. Note, while we can proportionally scale this range, final results in the model will not change, but translating it to the left or right, e.g. to [60,85], does alter final results. That said the ultimate skills vector for the ranked teams is given by $S = [90, 88, 84.8, 82.7, 80.3, 79.1, 77.5, 75.5, 75.1, 73.9, 73.1, 72.3, 71.5, 71, 70.6, 69.8, 69, 68.6, 67.4, 65]$. These are the values used to construct the geometric competition matrix. The shape of this curve indicates a linear model for skills may not work well – and it does not.

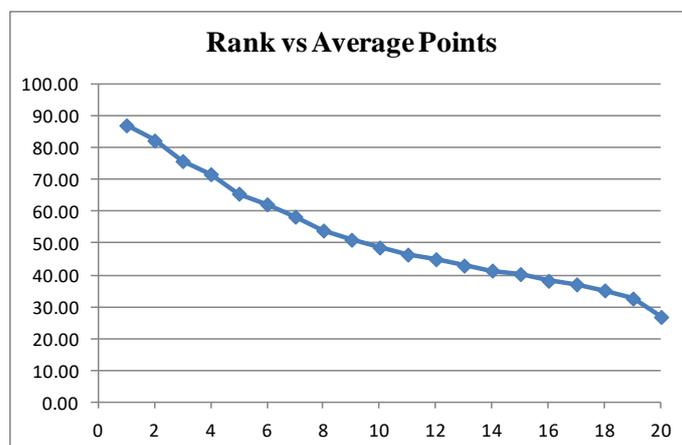


Figure 4.1 EPL Rank vs Average Points

If there were no ties, phenomenological modeling for soccer would be relatively simple, or a simplistic model could generate excellent results. However, there are ties. Looking at the average data over the past nine EPL seasons we see ties following a quadratic path as in Fig. 4.2. It illustrates a drift upwards as the rank of the team decreases in overall skills. Note the two peaks, one occurring for Team 11 and the other for Team 16. In fact, both of these peaks are significantly higher from the mean using a z-test at the 0.01 level. At least for Team 16 and sometimes for Team 17, there may well be the prospect of regulation that could account for this. By taking the maximum number of ties for teams 16 and 17, a rather nonstandard statistic, the significance from the z-test confirmed at every standard level. This particular observation must be built into any model, as has been in ours. The correlation of ties with a linear regression model, while revealing the upward drift gives the rather low correlation of $R^2 = 0.1659$; thus the linear model for ties should be discarded.

What is also remarkable about this graph is that it was generated from averages over multiple seasons. To account for the concave shape of the curve, called the **tie signature** of the league, we argue that the more evenly balanced the teams are the more likely there is a tie. The teams at either end of the rankings play fewer games with evenly matched teams, indicating the lower outcome of fewer ties. Indeed, for the top ranked team, there were on average fewer ties at the 0.01 level. This particular observation must be built into any model. The correlation of ties with a linear regression gives the rather low correlation of $R^2 = 0.1659$, thus a linear model for ties should be discarded.

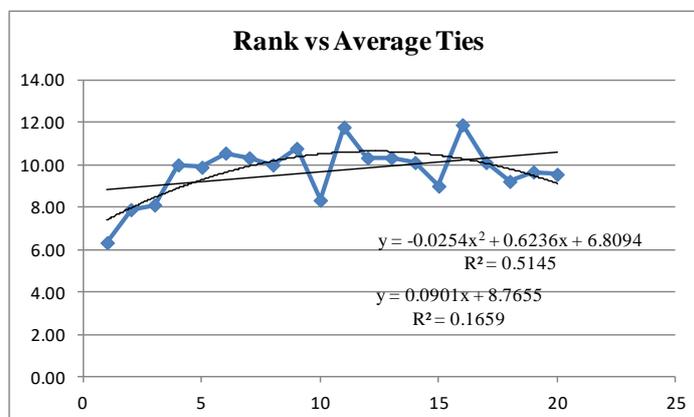


Figure 4.2 EPS Ties vs Rank

VI. The model

With the skills levels determined, the competition matrix computed, and the number of events per game decided (about 77), we need now to construct the model for win, tie, or loss records for all teams. This is achieved through a multiple trials type construction using a combination of integer and continuous arithmetic, together with a middle outcome range, wherein a tie is determined. Without venturing overly deep into details, we give the following model description.

1. For Team i and Team j , with skills $S_i \geq S_j$ we have a probability of relative success, c_{ij} . Normally, $0.5 \leq c_{ij} \leq 1$.
2. The next steps now pertain to a particular event within a match. For each event
 - a. Compute uniformly distributed random numbers in $[0,1]$.
 - b. For each number x determine whether it is less or greater than c_{ij}
 - c. If less than c_{ij} add one to success counter; if greater add one to the non-success counter. Actually, we only need the success counter, as the other is complementary with respect to the number of events.
3. Our use of random values requires twelve significant places of accuracy. With a greater granularity of such numbers, different outcomes are predicted. With only four digits, vastly different outcomes are predicted, in so far as statistics are concerned.
4. When all events have been exhausted, we award a win to Team i if the counter exceeds a certain value and a win to Team j if the counter is greater than another value. If the number of wins is between the two values, a tie to both teams is awarded. These four values are termed **event horizons**.

Moreover, event horizons are needed to differentiate home and away games, given that team skills are constant. (See Section 7 for further comments on this.) It is one thing to model ties and even wins based on historical data. Predicted outcomes will be as expected. It is another to generate fundamental assumptions which apply to all matches irrespective of such data beyond an assessment of skills. Many other simulations select probability ranges for the random numbers (in $[0,1]$) to describe win, tie, or loss, making predictions easier but less natural. Depending on whether the game is away or at home, the bandwidths and positions are differentiated. The skills remain the same, as is certainly the case. Simply put, when the numbers of significant events are approximately equal according to home or away game bandwidths, there is a higher likelihood of a tie. As a note, for a sport without ties, the bandwidths are collapsed to zero, the event horizons adjusted for the sport and league.

As is evident, we have converted home advantage to positions and widths of bandwidths. Now, with this model, we have yet to distinguish between relative skills. In generating this model, it was determined there is no relative combination of skills, number of events, and bandwidth with which a typical season could be modeled. To this end we modify, according to the number of relative skills, the numbers of significant events.

This is achieved only upon c_{ij} using the factor $\min\left(\frac{c_{ij}}{c_{ji}}, \frac{c_{ji}}{c_{ij}}\right)$, and the model is complete. It is the nature of

the tie that makes this variation of a multiple trials experiment both relevant and reasonable.

In a sense, the introduction of significant events is similar to assigning to each term of the competition matrix as probability density with given mean c_{ij} . The nature of the density requires some further research. This is alternative approach not addressed here.

The expression for this model could possibly take the more traditional form as

$$w_{ij} = f(c_{ij}, g_{ij}, E_{w_{ij}}) \quad \text{and} \quad t_{ij} = g(c_{ij}, h_{ij}, E_{t_{ij}})$$

where these quantities give the expected number of wins and ties for matches between Team i and Team j . Here g_{ij} and h_{ij} are vectors codifying additional parameters about the teams, and the terms $E_{w_{ij}}$ and $E_{t_{ij}}$ are stochastic factors. The functions f and g are appropriate to the model. However the model is designed, in simulations there is always a randomly generated number or many used to determine the stochastic factors. The steps 1, 2, and 3 above detail the construction of the functions. Often, the stochastic terms are introduced as additive effects to an expectation function. Thus for example, we commonly see expressions such as $w_{ij} = f(c_{ij}, g_{ij}) + E_{w_{ij}}$ and $t_{ij} = g(c_{ij}, h_{ij}) + E_{t_{ij}}$

As discussed previously, all these factors are relevant, reasonable, and significant. As is apparent, it is the nature of a tie that complicates the model and the computations. Let us note that because of the large number of events it would have been possible to use the normal approximations to the binomial distributions created and bi-pass the events scenario. However, the combination of integer and continuous arithmetic creates an ever so

slight additional randomness, upon which the normal season turns. Another sport to which this model applies is tennis. Each game in a set is a sequence of multiple trials, with “home” and “away” governed by which player has the serve. When a set reaches a tie at 6-6, there is a continuation, after which a slightly different model is employed in order the winner be determined. The key notion here is that winning a tennis match is not simply determined by the flip of an (unfair) coin, but is a sequence of a significant number of repeated trials (each point, and then each set) during which ties must be reckoned.

Baseball is a variant, also denying ties, enjoys different considerations. Models for both sports are under development. In both, an application of prospect theory seems to help predict outcomes. In football as discussed here, there must be in the case of major games the additional requirement of no tie, and this necessitates the imposition of further conditions, specifically that scores result from winning a sequence of significant events.

VII. Results

In this section, we describe the winning records, the collection of ties, and final rankings compared to initial rankings, and their related statistics. With the parameters for skills, the appropriate completion matrix, the bands for ties (home and away), we ran the program (written in Maple) for 100 seasons several times. The average number of points over these runs was 87.52 with a standard deviation of 2.15. This is comparable to actual records over the 06-07 to 14-15 seasons which were respectively 87.33 and 3.02. The implication is the model is slightly more stable in terms of variance, though we are comparing 100 seasons with just nine. This breaks down to the league winner having a average of 26.98 wins and 6.58 ties in the model. This compares with the actual data of 26.89 wins with and 6.33 ties.

In computing a χ^2 test with the actual average season W-D-L records over nine seasons under consideration with the average simulated season, we are concluded to reject the independence null hypothesis between the seasons at all levels, even the very smallest. This is an indicator that the model functions properly.

Table 7.1 Ties by Season and Original Rank

Season	Original Rank																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	3	6	5	6	8	2	4	5	8	4	4	5	5	5	5	5	2	7	7	4
2	11	6	10	11	6	9	15	9	12	6	10	12	8	5	10	11	13	14	8	12
3	11	4	9	9	9	7	13	12	15	6	10	11	10	10	6	14	6	11	11	8
4	8	9	9	10	7	6	15	6	8	12	9	9	10	12	8	7	10	10	10	11
5	6	9	8	11	10	9	11	11	15	7	8	6	12	6	9	10	8	9	5	8
6	8	9	13	11	5	10	9	8	5	13	7	13	7	5	11	12	8	5	11	6
7	8	7	10	9	10	6	6	9	7	7	6	12	4	11	6	10	9	9	5	5
8	5	4	8	10	6	10	10	7	12	6	7	10	11	8	8	11	10	7	6	6
9	7	6	6	8	7	11	8	10	10	10	9	10	8	7	10	12	4	10	8	7
10	7	7	5	5	7	9	6	9	7	13	7	9	14	11	4	6	7	17	8	10
11	7	10	7	12	8	9	11	9	9	13	11	10	17	12	11	7	11	12	10	6
12	8	6	8	14	8	10	12	9	5	14	8	9	15	12	11	11	11	6	11	6
13	2	4	4	7	8	11	9	10	5	6	7	7	9	3	11	7	5	6	7	2
14	6	8	12	12	7	8	9	7	12	10	7	11	11	11	10	10	8	7	7	9
15	7	12	12	8	12	12	10	9	11	11	12	8	13	11	8	18	9	10	6	5

Next, we come to the problem of ties, and just how well the model mimics the actual data. In Table 7.1 we show numbers of ties where for just the first fifteen seasons by the teams in order of their original rank. (The full table has one hundred lines.) However, all conclusions are made from the full table.

The average number of ties for the given rank over all seasons is [7.08, 7.81, 8.3, 9.21, 8.66, 8.88, 9.29, 9.19, 9.35, 9.57, 9.46, 9.36, 9.49, 9.13, 9.05, 8.94, 9.23, 9.19, 8.64, 7.81]. This information is illustrated in Fig. 7.1. It exhibits the same type of quadratic trend as we saw in Fig. 4.2 for the actual data. This confirms the notion about the model that the more equal teams are, the more likely a tie.

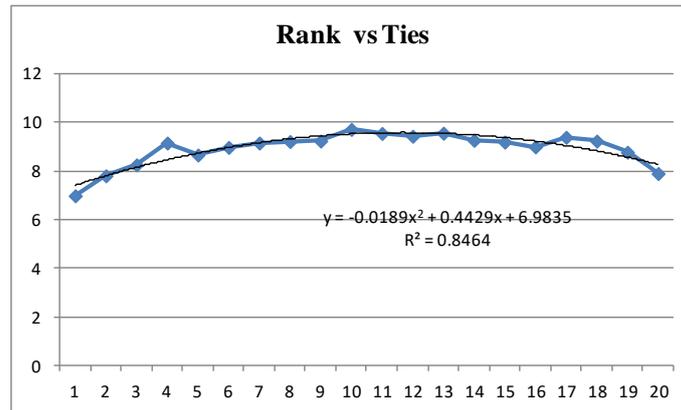


Figure 7.1 Average Ties by Original Rank

The one unsettling aspect is that the average number of ties is a bit lower for the model than actual data, 8.91 vs. 9.71. The special peaks notices for actual data are not there as well. However, our model did not account for that most interesting phenomenon. It may be of some interest to note the winning league team had overall the lowest number of ties, about seven.

Who wins and who loses? The final question we take up is who wins the championship? Is it the first ranked team all the time, or just a lot of the time? And which teams are relegated, at least most of the time? Again, for the actual data we know who has won each year, but for the simulated model, we need to count victories for each ranked team, and what place they took in the final standing – in total. Again, we dispense with the displaying the full table, allowing only the final table of rankings. So, we ask the question how many times over all the seasons does a ranked three team win the championship? Or how many of the lower three ranked teams are destined to be relegated at the end of the season. While Table 7.2 is given in frequencies, they automatically convert to probabilities, simply dividing by 100. For example, it is evident that the first ranked team placed third nine times over the 100 seasons, and the last ranked team never placed higher than 10th overall (twice). There is a great difference in the probabilities of outcome for teams in various ranks. Reality suggests skill levels of teams change through out the season. This will directly affect the competition matrix. Yet, how to make these changes is subject to conjecture. In the current model, the competition matrix can be changed at any time during the season based presumably on skills changes. Perhaps a correction should be introduced at the mid season and the three quarter season mark, times when teams can be reviewed on the one hand and get very serious about the season conclusion or simply begin their surrender.

The law of constancy: From this table of frequencies we can compute probabilities over many possibilities. For example, the probability that a top-half ranked team remain top-half ranked is 0.86, and those in the bottom half of rankings remain so is also 0.86. In fact this must be so by construction. The probability that top three ranked teams remain so is 0.81, while the probability that the bottom three teams remain so is only 0.57. Indeed, it seems easier for lower ranked teams to break out. Top teams on the other hand enjoy the constancy of their final position remarkably well. Lower ranked teams on the other hand suffer a relative constancy toward the bottom. Midrange teams have their own law of constancy, being somewhat constrained to remain in the

Table 7.2 Numbers of Final vs Original Ranks

Final Rank	Original Rank																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	61	18	11	8	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	27	33	25	11	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
3	9	26	33	13	13	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0
4	2	16	16	36	9	11	6	2	2	0	0	0	0	0	0	0	0	0	0	0
5	0	3	9	10	36	16	16	5	2	2	1	0	0	0	0	0	0	0	0	0
6	0	2	3	16	17	19	13	15	5	2	5	2	1	0	0	0	0	0	0	0
7	1	2	1	3	12	23	17	18	11	7	1	2	1	0	1	0	0	0	0	0
8	0	0	0	0	5	9	19	17	10	12	6	6	2	7	2	4	1	0	0	0
9	0	0	0	2	5	9	7	7	14	9	12	6	12	3	5	3	2	4	0	0
10	0	0	1	1	0	4	7	8	14	14	10	7	7	11	3	2	6	1	2	2
11	0	0	0	0	1	3	3	9	11	7	10	12	12	12	7	4	2	4	3	0
12	0	0	1	0	1	0	6	8	7	16	9	12	7	7	10	10	3	3	0	0
13	0	0	0	0	0	0	0	4	2	6	9	18	12	8	7	11	10	5	7	1
14	0	0	0	0	0	0	4	2	4	6	6	9	6	7	19	10	7	9	9	2
15	0	0	0	0	0	0	0	0	4	5	9	7	12	12	9	11	10	8	9	4
16	0	0	0	0	0	0	0	2	6	4	12	6	5	10	10	9	7	16	9	4
17	0	0	0	0	0	0	0	0	4	3	4	8	8	11	10	12	12	14	8	6
18	0	0	0	0	0	0	2	1	5	3	3	6	7	7	10	20	16	11	9	
19	0	0	0	0	0	0	0	0	1	2	3	1	8	3	8	9	16	12	19	18
20	0	0	0	0	0	0	0	0	0	0	1	1	2	2	5	4	8	23	54	

middle. Yet, the exceptions make the sports enterprise interesting and sustaining. For the very bottom ranked teams, it can be argued that relegation merely serves to supply new fodder for the established higher quality teams. This is not to say relegation is without merit. In Table 7.3 we consider only a few possibilities, including those just mentioned, of league outcomes and their contingent probabilities.

Table 7.3 Probabilities of Select Outcomes

Probability that	Value
All three bottom teams are relegated	0.57
Both two bottom teams are relegated	0.67
The bottom team is relegated	0.81
Teams originally ranked 8-12 remain so	0.51
Teams originally ranked 6-14 remain so	0.84
Team ranked team 6th drops at least one rank	0.43
Teams ranks top three remain so ranked	0.81
Teams not ranked first win the championship	0.39

Clearly, teams originally ranked in the broad middle range tend to stay there with a high probability. Naturally, as these frequencies are computed from a simulation, the second decimal point above is suspect as it applies to the actual football world. From this table, the interested reader can compute probabilities for any virtually any hypothetical. One item apparent is that bottom-ranked teams stay there. If not for the loyalty of fans, it is likely the EPL would have to resort to a draft to give more competitive balance to the teams, acceding to the uncertainty of outcomes hypothesis. This can be modeled by simply changing the range on the skills vector. For example, changing the skills range from [65,90] as we used to [70,90], with all other factors constant, a considerable change in final standings would result.

There is a great difference between probabilities of outcome for teams in various ranks. This assumes skill levels of teams change during the season, and this directly affects the competition matrix. Yet, how to make these changes is a matter of conjecture, and conjectures may vary widely. Nonetheless, in the current model, the competition matrix can be changed at any time during the season based presumably on skills changes. Perhaps a correction should be introduced at the mid season and the three-quarter season mark, times when teams review past performance and get very serious about winning and the season's conclusion.

Other data could be reported such as to address the home-away W-D-L records over the 100 seasons between Team 2 and Team 12, with implications toward creating probabilities of a win, tie, or loss on game day. Finally note all row and column sums of this frequency matrix sum to one hundred. Therefore, dividing all entries by 100, the matrix becomes a doubly stochastic matrix and in fact a transition matrix. This permits the consideration of probabilities of multiple seasons based on the previous seasons. It is not entirely clear how to interpret these possibilities, but they are available for consideration.

VIII. Conclusions

The results of the previous section indicate the present model produces remarkably accurate estimates of actual EPL seasons. It even produces seasons with relatively low point totals, similar to actual but relatively uncommon EPL seasons. Our competitive balance model seems to reveal interesting results, particularly about final placement rankings. What is significant is that a generalized model is not possible applying across leagues or sports, without league specific data. Namely, it is important to have the point and ties profile signatures for the league considered. As well, it is important, in this model event profiles and critically important an estimate of the numbers of significant events during a single match. The event horizons were selected to accommodate the tie profile for the home-away exigencies. The number of events was a guess. But all the factors, skills, event horizons, events, games per season, home-away factors all co-mingle in highly nonlinear relationships. Using the suitable competition matrix is also important.

The approach taken was foundational using parameters that are reasonable, (probably) computable, based on real data, and parsimonious. However, as we've noted along the way, the model has flaws directly connected with actual season play. Here are a few questions, some reflective on earlier work, and comments for future work.

1. What is the probability that a team looking at relegation at the $\frac{3}{4}$ season point actually are relegated. This could affect skills levels and hence probability levels.
2. Should team records at the half season be marked so to recalibrate the relative skills for the remainder of the season?
3. Is it important to consider seasonal trends, perhaps moving averages over multiple seasons?
4. As most scores involve anywhere from one to several events won in sequence, should the model include the presence of consecutive runs of winning events? This would be a precursor to estimating numbers of game goals [16].
5. Should adjustment of skills of game-day be based on betting odds [23]? Possibly, but the bookmakers are not necessarily that accurate.

For some of these questions, an enhanced model can be devised, inserted and better results, certainly more stable results obtained. However, doing this creates a risk of seeming to “cook” the books. There is also the matter of making reasonable assumptions about the relative importance of a host of variables particularly when including game-day considerations. Historical factors may yield more reliable results. Learning methods from AI such as random forests may help to sort the relative importance of factors. The code, in Maple, is freely available upon request.

References

- [1] England Football Results Betting Odds — Premiership Results & Betting Odds. Web. 11 Dec. 2014. <http://football-data.co.uk/englandm.php>
- [2] Rottenberg, S., (1956). The baseball players' labor market. *Journal of Political Economy* 64, 242-258.
- [3] Forrest, D., Simmons, R., (2002). Outcome uncertainty and attendance demand in sport: The case of English soccer. *Journal of the Royal Statistical Society, Series D (The Statistician)* 51, 229-241.
- [4] Noll, R.G., (1988). Professional Basketball, Studies in Industrial Economics Paper No. 144. Stanford University, Stanford, CA.
- [5] Scully, G.W., (1989). *The Business of Major League Baseball*. (University of Chicago Press, Chicago, IL).
- [6] Quirk, J., Fort, R.D., (1992). *Pay Dirt: The Business of Professional Team Sports*. (Princeton University Press, Princeton, NJ).
- [7] Fort, R., Quirk, J., (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature* 33, 1265-1299.
- [8] Schmidt, M.B., Berri, D.J., (2001). Competitive balance and attendance: The case of Major League Baseball. *Journal of Sports Economics* 2, 145-167.
- [9] Humphreys, B.R., 2002. Alternative measures of competitive balance in sports leagues. *Journal of Sports Economics* 3, 133-148.
- [10] Dobson, S., Goddard, J., (2010). Optimizing strategic behaviour in a dynamic setting in professional team sports. *European Journal of Operational Research* 205, 661-669.
- [11] Brizzi, M., (2002). A class of indices of equality of a sport championship: Definition, properties and inference. In: Mrvar, A., Owen, P. Dorian, and King, Nicholas, (2015) *Competitive Balance Measures In Sports Leagues: The Effects of Variation in Season Length*, *Economic Inquiry*, (ISSN 0095-2583), Vol. 53, No. 1, 731-744.
- [12] Clarke, S.R., (1993). Computer forecasting of Australian rules football for a daily newspaper. *Journal of the Operational Research Society* 44, 753-759
- [14] Koning, R.H., Koolhaas, M., Renes, G., Ridder, G., (2003). A simulation model for football championships. *European Journal of Operational Research* 148, 268-276.
- [15] Forrest, D., Beaumont, J., Goddard, J., Simmons, R., (2005). Home advantage and the debate about competitive balance in professional sports leagues. *Journal of Sports Sciences* 23, 439-445.
- [16] Goddard, John, (2005), “Regression models for forecasting goals and match results in association football.” *International Journal of forecasting* 21.2, 331-340.
- [17] Joseph, A. E. Fenton, & M. Neil., 2006, Predicting football results using Bayesian nets and other machine learning techniques. *Knowledge-Based Systems* 19.7, 544-553.
- [18] Bradley, R.A., Terry, M.E., (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika* 39, 324-345.
- [19] Breiman, Leo, (1968) *Probability*, Addison-Wesley, Boston.
- [20] Kahneman, Daniel; Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk" (PDF). *Econometrica* 47 (2): 263. doi:10.2307/1914185. ISSN 0012-9682.
- [21] Tversky, Amos; Kahneman, Daniel (1992). "Advances in prospect theory: Cumulative representation of uncertainty". *Journal of Risk and Uncertainty* 5 (4): 297–323. doi:10.1007/BF00122574. ISSN 0895-5646.
- [22] Burgos, Albert, (2004) Guessing and Gambling”, *Economics Bulletin*, (2004) Vol. 4, No. 4, pp. 1-10 (<http://www.economicsbulletin.com/2004/volume4/EB-04D80001A.pdf>)
- [23] Crowder, M., Dixon, M., Ledford, A. and Robinson, M., (2002), “Dynamic modelling and prediction of English Football League matches for betting”. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 51: 157168. doi: 10.1111/1467-9884.00308.