Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

WARRIORS OR BULLS: INTRODUCING RETROACTIVE GAMBLING LINE SIMULATION

R. Alan Bowman

School of Business
Clarkson University
Capital Region Campus
80 Nott Terrace
Schenectady, NY 12308, USA

Thomas Ashman

Behavioral Sciences Collegium
Eckerd College
4200 54th Avenue S
St. Petersburg, FL 33711, USA

James Lambrinos

School of Business
Clarkson University
Capital Region Campus
80 Nott Terrace
Schenectady, NY 12308, USA

ABSTRACT

Gambling lines provide a rich source of data for Monte Carlo simulations of sports seasons. Since data for complete seasons are available after the seasons are completed, we refer to such simulations as retroactive gambling line simulations. We use the motivating and expository example of comparing the performances of the 2015-16 Golden State Warriors (who won a record 73 games) and the 1995-96 Chicago Bulls (who held the previous record of 72 wins) and provide a detailed treatment of this application. The more general applicability of the approach is examined by describing the flexibility in building the models and relating this flexibility to a few specific examples. We hope this paper will seed further explorations of retroactive gambling line simulation (and settle the Warriors vs. Bulls comparison along the way).

1 INTRODUCTION

The Golden State Warriors won 73 games in the 2015-16 NBA season, breaking the record of 72 wins previously set by the 1995-96 Chicago Bulls. The pursuit of the record was one of the most widely covered sports stories in recent times and a common topic of discussion was to compare the achievements of those two teams. Whose season was truly more impressive considering the teams they played, home advantages, injuries, travel, rest, and other schedule factors, etc.? Having used gambling line data to assess competitive balance (Bowman, et al 2013a and 2013b) and to evaluate the impressiveness of winning streaks (Bowman, Ashman, and Lambrinos 2015), we recognized the potential power of gambling line data to effectively compare the Warriors' and Bulls' achievements in an objective manner since each game's gambling line should consider all factors affecting the likelihood of each team winning. The two most common forms of gambling lines are point spreads and money lines. A point spread tells how many points one team is favored to beat the other team by and to win a bet on the favorite would require that team to win by more than the point spread. A positive money line (underdog) tells how much a gambler would win per $100 bet and a negative money line (favorite) tells how much a gambler would have to bet to win $100. Both forms can be used to derive game probabilities for use in Monte Carlo simulations.

In this paper, we present both analytical and Monte Carlo simulation approaches for comparing the performances. The results indicate that the simulation approach is significantly more accurate for assessing achievements "in the extreme tails" of a distribution, which we are doing here. In developing the framework for the Monte Carlo simulation approach to answer this question (and the full flexibility possible within that framework), we believe we have identified an approach that can be used to analyze a wide variety of issues either more accurately than analytical approaches can, or when analytical approaches are not possible. In this paper, we present a full analysis of the Warriors' and Bulls' achievements for readers interested specifically in that comparison. More importantly, in our opinion, we also use this analysis for exemplification and expository purposes. In that sense, our main goal is to introduce the simulation approach, which we call retroactive gambling line simulation, to the literature and provide initial insights into the potential variety of applications. We note that this methodology can be applied to address many different questions in a particular sport and can also be used to address these questions for any sport that has active gambling lines.

In the next section, we review the literature to place retroactive gambling line simulation in context. In sections 3 and 4, we then describe an analytical approach for using gambling lines to compare the Warriors and Bulls and follow this comparison through to both introduce and exemplify retroactive gambling line simulation. We discuss the general capabilities of retroactive gambling line simulation and a few specific potential applications. In section 5, we present both our analytical and simulation results for the Warriors/Bulls comparison. We are interested in both what the results say about the comparison and also how the analytical and simulation assessments compare. In section 6, we analyze the likelihood of an event that very recently occurred for the first time in sports history using retroactive gambling line simulation as a relatively simple example of what we believe to be a large class of issues that retroactive gambling line simulation can effectively address that are not easily amenable to analytical approaches. Finally, in section 7, we draw some conclusions about retroactive gambling line simulation and about the Warriors and the Bulls as well.

2 LITERATURE REVIEW

There has been considerable, if not widespread, use of Monte Carlo simulations in sports-related research. We divide the research into two broad areas of inquiry. The first area is the measurement of either technical efficiency or the performance of individual players or teams. For example, Rimler, Seong, and Yi (2009) use Monte Carlo techniques to estimate the technical efficiency of National Collegiate Athletic Association (NCAA) men's basketball teams. Koop (2002) uses Monte Carlo simulation to compare the performance of Major League Baseball (MLB) hitters. Lenten (2016) examines incentives in the NBA and MLB for teams to underperform due to reverse-order drafting. Cote, Macdonald, Baker and Abernethy (2006) use Monte Carlo simulations to assess whether birthplace and/or birthdate affect the probability of becoming a professional athlete.

The second area of sports research where Monte Carlo simulation has been used is the prediction of the outcome of individual sports contests or groups of such contests. Our paper falls in this broad area. One of the earliest models to use Monte Carlo simulation to assess individual game probabilities was Glickman and Stern (1998). They devised a model that outperformed Las Vegas point spreads in predicting the margin of victory in National Football League (NFL) games. They used data from the 1988 through 1993 seasons, and incorporated the effects of injuries, season-to-season changes in personnel, and home-field advantage. The model they constructed outperformed closing Las Vegas point spreads over the last 110 games of the 1993 NFL season. They used 300 samples of parameters from the posterior distribution and simulations from the predicted data to diagnose invalid assumptions or lack of fit of their model. The simulations suggested that variance of the actual data was consistent with what was expected under the model.

Strumbelj and Vracar (2012) constructed a model that simulates the play-by-play progression of individual basketball games. They estimated transition (each "transition" refers to an event during the game that corresponds to a team retaining or acquiring possession) probabilities for the games from play-by-play data for the 2007-2008 and 2008-2009 seasons. From this transition matrix, Strumbelj and Vracar estimated win probabilities using 10,000 simulated matches.

Gupta (2015) used Monte Carlo simulation to predict "bracket" performance for bettors in the NCAA Division I Men's Basketball Tournament. He constructed a rating system using a dual-proportion model based on game data from the NCAA season prior to the tournament. He examined six NCAA seasons and tournaments, so his predictions were not validated by only one season. He did not use gambling odds for his rating system (as we do in this paper). Using his rating system, Gupta determined the probability that predicted the champion first and then worked backwards to predict the winners in the round of 64. Each game's marginal probability consisted of the probability of winning the game and the conditional probability of reaching the game. The bracket so constructed was the one believed to have the highest expected value of winning the bracket competition. Gupta then used 10,000 Monte Carlo simulations for various pool sizes to evaluate the effectiveness of twenty different bracket strategies for winning NCAA tournament pools in 2013 and 2014. Nineteen of these strategies involved a modification to the "All Favorites" bracket (e.g., the champion from "All Favorites" was replaced with the second-highest-probability champion). While the "All Favorites" (i.e., highest expected value) strategy won a large proportion of the simulations, for each pool size (11, 21, 41, 101, and 1001), there was a modification to the "All Favorites" strategy that won a higher proportion of pools in both years.

Barry and Hartigan (1993) predicted divisional winners in the National League using data from the 1991 season. They developed a choice model which used data from the early games of the season and estimated the impact of changes in strength of the teams during the remainder of the season. A key parameter in their model was the proportional change in strength, which they estimated using Markov chain sampling. They predicted divisional winners with 2,500 simulations at various points of the season.

Our approach differs from any of the previous uses of Monte Carlo simulation in two fundamental ways. First, we use gambling line data. Second, we simulate seasons retroactively. Our approach is not meant to predict what might happen in a game or season or tournament that is coming up. Instead, our approach can look back at a season (or multiple seasons) and estimate probabilities that certain events might have happened (what did happened is viewed as simply one possible sample path). As we will see, the gambling lines can be broken into components and then the components can be used in the simulations in various combinations depending on what the researcher is investigating. We are not limited to simply simulating the seasons according to the probabilities inferred directly from the gambling lines. For example, we can alter the schedule and we can consider specific game factors as they occurred or in other ways befitting the analysis.

The chief advantage of using gambling lines is that they presumably take everything into account that might affect the outcome of a game (or else it would be possible for sophisticated gamblers to exploit factors that were not taken into account). This presumption is consistent with the "balanced book" model of optimal sports book behavior. That is, the balanced book model assumes that the sports book sets the point spread or money line so that money flows occurring from bets are equal for both teams. In this way, the sports book derives a riskless profit from the "vigorish" or commission inherent in the payoffs. Levitt (2004) challenged the assumption that sports books optimize their profits by equating money flows on the teams involved in a game. He observes that sports books have better information than most bettors and can therefore profit by setting a spread that takes advantage of naïve or uninformed bettors. There has been empirical evidence (e.g., Paul and Weinbach (2007, 2008)) to support Levitt's contention. Nevertheless, Levitt indicated that the variance from "true" spreads is likely small because of the risk that sports books would bear from informed bettors taking advantage of persistent mispricing. This ability of the gambling line markets to efficiently reflect all factors known at the time of the match is fundamental to our approach but also means that our analyses must be retrospective in nature. Of course, it is typical that the whole point of retrospective analyses is to shed light on the future and that is true here as well.

3 A PRELIMINARY (NAIVE) APPROACH

3.1 The Approach

Bowman, Ashman, and Lambrinos (2013b) utilized gambling line information to assess competitive balance in major league baseball. One measure they suggested (and computed) involved converting money line information into game probabilities and using the following formulas to compute the expected value and standard deviation of the number of wins for each team in the league. We have altered the formulas only to reflect that NBA teams play 82 games whereas MLB teams play 162.

$E(W_{ik}) = \sum_{j=1}^{82} p_{ijk}$, and

$SD(W_{ik}) = \sqrt{\sum_{j=1}^{82} p_{ijk}\,(1 - p_{ijk})}$, where

$E(W_{ik})$ = the expected value of the number of wins of team i in season k,
$SD(W_{ik})$ = the standard deviation of the number of wins of team i in season k, and
p_{ijk} = the probability that team i wins game j in season k.

Note that the main form of gambling lines in MLB (and the NHL) is money lines whereas it is point spreads for the NBA (and the NFL). Bowman, Ashman, and Lambrinos (2015) converted the spreads to game probabilities in order to compare the 4 major sports leagues and we followed their approach for the NBA to produce the results in this paper.

Since we are summing a fairly large number of independent random variables (they are independent because the gambling lines, if efficient, already take into account all the dependencies), the number of wins should be approximately normally distributed. We can use these results to compute the probability that the Warriors would win 73 or more games and the probability the Bulls would win 72 or more games. If we do so, the results are .0306 and .0493, respectively. Perhaps we could conclude that the Warriors' season was more impressive based on this.

3.2 Issues

There are two issues, however, that make this assessment dubious at best:

1. The better a team is, the higher its game win probabilities are, hence the higher its chances of winning at least a certain number of games, hence the less impressive any particular number of wins would be. Doing this really measures performance against expectation, which might be interesting as well, but is not what most people would mean by impressive.
2. If we are dealing with small probabilities (the tails of the distribution), the normal distribution may not be an accurate approximation.

4 IMPROVING THE APPROACH

4.1 Removing the Effect of the Team Itself

Bowman, Lambrinos and Ashman (2013a) provide a means for addressing the first issue via the model of spreads they presented, which they used to explore competitive balance more deeply. We review the model below:

$S_{jk} = N_{ijk} – H$, where
S_{ijk} = the spread when team i is playing at home against team j in game k,
N_{ijk} = the "neutral" spread – the spread adjusted by the average home advantage, and
H = the average home advantage (averaged across all games in the season).
They further broke down N_{ijk} as follows:
$N_{ijk} = -(R_i – R_j + \varepsilon_{ijk})$, where
R_i = the rating of team i (how much team i would be favored against the average league team in a neutral site game based on the entire season – note that this means the average of all the ratings is 0),
ε_{ijk} = the factors specific to team i playing team j in game k.

They suggested estimating the team ratings using the following regression model with the ε_{ijk} terms being the error terms in the model:

$S_{h(k)a(k)k} = \beta_0 + \sum_{i=1}^{m-1} \beta_i X_i + \varepsilon_{h(k)a(k)k}$, where

$$X_i = \begin{cases} -1 \text{ if } h(k) = i \text{ (team i is the home team in game k)} \\ 1 \text{ if } a(k) = i \text{ (team i is the away team in game k)} \\ 0 \text{ otherwise} \end{cases}$$

and m = the number of teams in the league.

Note that one of the teams (team m) is left out of the model and that the coefficient estimates are relative to that team. To make the ratings estimates invariant to the team left out (and relative to the average team as required by the definition of the ratings terms), the coefficients are converted to the ratings as follows:
$\widehat{R_i} = \widehat{\beta_i} - \frac{1}{m} \sum_{j=1}^{m-1} \widehat{\beta_j}$, i = 1 to m-1, and
$$\widehat{R_m} = -\frac{1}{m} \sum_{j=1}^{m-1} \widehat{\beta_j} \ .$$
The key point for our purposes is that this model allows us to substitute in any team's rating R (or any rating R at all for that matter – it doesn't need to correspond to any actual team) for either team involved in a game to compute what the spread would be if the one team's rating was replaced with R. This would leave the home court advantage H and the game specific effects ε_{ijk} as they were. If we replaced the Warriors' (Bulls') rating with R for every game they played in the 2015-16 (1995-96) season, we could use the expected value and standard deviation of the number of wins and the normal approximation to estimate the probability that a team with rating R would win 73 (72) or more games playing the Warriors' (Bulls) schedule in 2015-16 (1995-96). We could do this for any value of R (the key being to use the same R to assess both teams and thus not have a team's rating influence its impressiveness measure). This addresses the first issue with the preliminary approach.

4.2 Introducing Retroactive Gambling Line Simulation

To address the second issue, we introduce the methodology we will refer to as retroactive gambling line simulation. The most fundamental type of retroactive gambling line simulation would be to simulate a selected number of replications of a season using the actual games played in that season and the game win probabilities from the gambling lines. This essentially views the actual outcome of the season as simply one possible sample path that the season could have taken. Conducting Monte Carlo simulations explores the entire population of sample paths from which many different summary measures can be computed.

One example, of course, is the motivating one for this paper that we have been following where the summary measure is the number of wins. It now makes sense, for example, to speak of the probability that the Warriors would win more than 73 games in the 2015-16 NBA season rather than just noting whether they did or not.

The fact that a vast array of season performance measures can be estimated using retroactive gambling line simulation already suggests an applicability well beyond our motivating example. We also note, however, that the gambling line model reviewed in section 3.1 breaks down the win probability (or spread, hence win probability once converted) into components relating to the teams involved, the home advantage, and the game specific factors. The game specific factors are very interesting and depend to some extent on the sport. They would include such things as when did each team play its previous game, how far they each had to travel from their previous games, who on each team was injured or ill, etc. – in general, anything affecting the win probability other than the baseline team differential and the home advantage. These components can be included or not included depending on what the researcher is investigating. For our motivating example, we clearly want to include the game specific factors as they occurred. If the Warriors faced teams with injured players less than the Bulls did, we want that to be reflected in the measures. However, as we have already noted, we will want to switch out the actual team whose performance we are evaluating and the component breakout allows us to do so.

The number of interesting potential investigations is large and each one could involve a different use of the components. Just to consider one other example, suppose we wanted to compare two possible scheduling algorithms in terms of how many times each team played each other team each season (each of the 4 major sports leagues does this differently and the NHL in particular has made some major alterations in its algorithm through the years). One interesting aspect of such decisions would be the effect they would have on the probability that each team makes the playoffs (to assess schedule equity). This could be addressed using retroactive gambling line simulations using the teams and home advantage from the various algorithms to be evaluated and randomly sampling from the specific game factors. In this and other applications, we think the ability to combine the components in different ways to investigate different issues greatly extends the applicability of retroactive gambling line simulation.

5 WARRIORS OR BULLS – THE RESULTS

5.1 Comparisons by Team Rating

Returning now to our motivating example, recall that the rating R for a team can be interpreted as the spread of the game if the team was playing the average team in the league on a neutral court and all game specific factors did not favor either team. We computed the probability that a team with rating R would win at least 73 (72) games playing the 2015-16 Warriors' (1995-96 Bulls') schedule using the analytical expected value and standard deviation with the normal approximation and also using retroactive gambling line simulation (100,000 replications) The results are shown in Table 1 for every value of R for which a team with that rating won at least 73 (72) games playing the 2015-16 Warriors' (1995-96 Bulls') schedule at least once in the 100,000 simulations all the way up to 10.8 (the highest rating of any team in our entire data set, which, by the way, was the 1996-97 Bulls). For brevity, results are only shown for ratings that actually occurred at least once in the data history (1990-91 through 2015-16).

Also shown is the historical proportion of teams with each rating. We can see in the table that, no matter what a team's rating, it would have had a better chance to win at least 72 games playing the Bulls' schedule than at least 73 games playing the Warriors' schedule. It is pretty clear that the Warriors' performance was more impressive.

100,000 simulations of the Warriors' and Bulls' seasons using the actual ratings of the Warriors and Bulls, respectively, resulted in the estimate of the probability that the Warriors would win 73 or more games as .0259 (compared to the analytical estimate of .0306) and the probability the Bulls would win 72

Table 1: Analytical and Simulation Estimates

Rating R	Historical Proportion of Teams with this Rating	Probability Team With Rating R Would Win 73 or More Games Playing Warriors 2015-16 Schedule		Probability Team With Rating R Would Win 72 or More Games Playing Bulls 1995-96 Schedule	
		Simulation	Normal Approx.	Simulation	Normal Approx.
4.8	0.0048	0.000000	0.000002	0.000010	0.000013
4.9	0.0012	0.000000	0.000003	0.000010	0.000017
5.0	0.0060	0.000000	0.000004	0.000010	0.000021
5.1	0.0048	0.000000	0.000005	0.000010	0.000026
5.2	0.0060	0.000000	0.000006	0.000010	0.000033
5.3	0.0036	0.000000	0.000008	0.000020	0.000042
5.4	0.0048	0.000000	0.000010	0.000020	0.000053
5.5	0.0048	0.000010	0.000013	0.000020	0.000067
5.6	0.0024	0.000010	0.000016	0.000040	0.000084
5.7	0.0024	0.000010	0.000020	0.000040	0.000105
5.8	0.0024	0.000010	0.000026	0.000040	0.000126
6.0	0.0060	0.000020	0.000041	0.000050	0.000196
6.1	0.0096	0.000020	0.000051	0.000070	0.000243
6.2	0.0096	0.000020	0.000062	0.000080	0.000289
6.3	0.0012	0.000020	0.000077	0.000100	0.000357
6.4	0.0060	0.000020	0.000096	0.000140	0.000437
6.5	0.0036	0.000020	0.000120	0.000190	0.000523
6.6	0.0048	0.000020	0.000142	0.000220	0.000638
6.7	0.0096	0.000040	0.000177	0.000320	0.000777
6.8	0.0024	0.000040	0.000219	0.000400	0.000918
6.9	0.0012	0.000070	0.000269	0.000500	0.001124
7.0	0.0012	0.000100	0.000321	0.000620	0.001327
7.1	0.0024	0.000140	0.000392	0.000790	0.001591
7.2	0.0036	0.000190	0.000468	0.001030	0.001920
7.3	0.0012	0.000230	0.000570	0.001180	0.002259
7.4	0.0012	0.000260	0.000700	0.001440	0.002694
7.5	0.0012	0.000300	0.000826	0.001750	0.003137
7.6	0.0012	0.000370	0.001001	0.002200	0.003729
7.7	0.0036	0.000440	0.001179	0.002750	0.004332
8.1	0.0012	0.001020	0.002368	0.005440	0.008124
8.3	0.0012	0.001620	0.003301	0.007620	0.011020
8.6	0.0012	0.002830	0.005301	0.012720	0.016823
9.4	0.0012	0.011300	0.017311	0.042000	0.046823
9.8	0.0012	0.020950	0.029318	0.070300	0.073697
10.8	0.0012	0.078060	0.091738	0.199880	0.197707

or more games as .0448 (compared to .0493 analytically). These simulation estimates have 95% confidence interval half-widths of .00050 and .00065 so the respective differences of .0047 and .0045 are both statistically and practically significant (differences of 15% and 10%). The differences become more significant as the probabilities get lower (the seasons more impressive). This can be more clearly seen in

Table 2. For this application, the increased accuracy of Monte Carlo simulation compared to invoking the Central Limit Theorem seems to be clearly worthwhile.

Table 2: Analytical vs. Simulation Estimates Summary Comparison

Team Rating Range	Average Probability of Winning 73 or more games vs. Warriors' schedule			Average Probability of Winning 72 or more games vs. Warriors' schedule		
	Simulation	Normal Approx.	% Error	Simulation	Normal Approx.	% Error
4.8 – 6.8	0.000013	0.000054	304	0.000088	0.000244	177
6.9 – 8.8	0.001135	0.002325	105	0.005400	0.007686	42
8.9 – 10.8	0.029587	0.038135	29	0.087469	0.090268	3

5.2 Summary Measure – Season Performance Impressiveness

Perhaps the single best measure of season performance impressiveness is the weighted (by historical proportion) average probability of winning the target number of games. In other words, we define the SPI (season performance impressiveness) of team T in season S as:

$SPI = \int_{R_S}^{R_L} P(R, X, T, S) f(R)\, dR$, where

R_S = smallest historical rating of any team,
R_L = largest historical rating of any team,
P(R, X, T, S) = probability that a team with rating R would win at least T games playing team X's schedule in season S (and T is the number of games team X actually won in season S), and
f(R) = the probability density function for R.

Defined in this way, the SPI can be interpreted as the probability that a randomly selected NBA team would win at least T games playing team X's schedule in season S.

We estimate f(R) using the historical proportion of teams observed with each rating (rounded to the nearest tenth) and P(R,X,T,S) either using the normal approximation or Monte Carlo simulation.
Using the estimates in Table 1, we get the estimates for SPI in Table 3.

Table 3: Season Performance Impressiveness (SPI)

Team	SPI (Simulation)	SPI (Normal Approximation)
2015-16 Warriors	.000166	.000197
1995-96 Bulls	.000440	.000495

As expected, since winning 73 games or more playing the Warriors schedule is more impressive than winning 72 or more games playing the Bulls schedule for every team rating R, the Warriors' SPI is smaller (more impressive) than the Bulls' SPI. The normal approximation is higher than the simulation estimates in both cases. This is also as expected since the normal approximation is fairly consistently higher in Table 1. Note, however, that we can compute SPI using any team's schedule and any specified number of wins – we don't have to use the number of wins the team actually had when playing the schedule. This enables us to compare the Warriors and Bulls more definitively. The SPI for winning 72 games playing the Warriors' schedule (using simulation) is .000332, compared to .00440 for winning 72 games playing the Bulls' schedule. Even if the Warriors had only won 72 games and tied the Bulls' record, their season would still have been more impressive.

For one more look at Warriors vs. Bulls and also one more look at the accuracy of the normal approximation, Table 4 shows the results for the Warriors and Bulls schedules and winning 62 or more and 75 or more games (the former representing an impressive but not at all rare achievement and the latter representing an unlikely but not impossible goal).

Table 4: Analytical and Simulation Estimates of SPI: Summary Comparison

Team's Schedule	Probability of Winning 62 or more games			Probability of Winning 75 or more games		
	Simulation	Normal Approx.	% Error	Simulation	Normal Approx.	% Error
2015-16 Warriors	0.0263	0.0264	<1	0.0000301	0.0000452	50
1995-96 Bulls	0.0299	0.0299	0	0.0000473	0.0000675	43

We can see that winning 62 or more vs. the Warriors' schedule is more impressive than winning 62 or more vs. the Bulls' schedule. The same is true for 75 or more games. This is actually true for all numbers of wins from 62 on up, which indicates the Warriors faced a tougher schedule (more on this later).

We can also see that the normal approximation is very accurate for the higher probability (62 or more wins) estimate but not nearly as accurate for the lower probability (75 or more wins) estimate. Since we are especially interested in the most impressive (low probability) estimates, this suggests that retroactive gambling simulation is an essential tool for this purpose.

In our data period (1990-91 through 2015-16), there were 26 teams that won 62 or more games. We computed SPI for each of these team's schedules and for each minimum number of wins from 62 up to all 82. For brevity, we do not report them here but they are available on request. We note that, in general, the SPI for each team's actual number of wins in its season was generally lower (hence more impressive) for larger numbers of wins. There were, however, notable exceptions. The Spurs 62-win season in 2013-14 had an SPI of .017457 whereas the Pistons 64-win season in 2005-6 had an SPI of .018880. This is an example where a team with two fewer wins was actually more impressive. A closer look at the numbers shows that the Spurs played teams with an average rating of .417 after adjusting the ratings to remove the Spurs (so that the average rating of all other teams was 0). This was the highest of any of the 26 teams with 62 or more wins. In comparison, the 2015-16 Warriors played teams with an average (adjusted) rating of -.046 and the average adjusted rating of the Bulls' opponents in 1995-96 was -.134. This means both teams played, on average, teams that were slightly worse than average (whereas the Spurs opponents were above average), with the Bulls' opponents being weaker than the Warriors.

We also computed the correlation of the game effects with the opponents' ratings. This correlation was .106 for the Warriors and .358 for the Bulls. The positive correlations indicate both teams got somewhat of a benefit by having the game specific factors (which average 0 for each team across a season) be more in their favor when they were playing better teams but the Bulls got more benefit from this than did the Warriors. In pretty much every aspect of the gambling line data, the Warriors' 73-win season was more impressive than the Bulls' 72-win season.

Just to satisfy sports fans' curiosity, we report the top 10 SPI's for the 1990-91 through 2015-16 seasons in Table 5.

Table 5: Top 10 Most Impressive Seasons (1990-91 – 2015-16)

Rank	Team	Season	Wins	SPI	Rank	Team	Season	Wins	SPI
1	Warriors	2015-16	73	.000166	6	Lakers	1999-00	67	.003746
2	Bulls	1995-96	72	.000440	7	Spurs	2015-16	67	.003910
3	Bulls	1996-97	69	.001731	8	Bulls	1991-92	67	.004268
4	Warriors	2014-15	67	.002755	9	Cavaliers	2008-09	66	.006594
5	Mavericks	2006-07	67	.003105	10	Celtics	2007-08	66	.008135

6 ASSESSING THE UNIQUENESS OF A FIRST TIME EVENT

The Warriors/Bulls comparison motivated the development of retroactive gambling line simulation, which we have seen enabled us to develop more accurate assessments of low probability events than could be obtained via analytical models. In that case, however, there was a relatively simple analytical model available. As this research was in progress, a sports event happened that had never happened before in the history of the 4 major sports. On December 31, 2016, the Columbus Blue Jackets played the Minnesota Wild in the first ever matchup of two teams with winning streaks of at least 12 games. What is the probability of such an event? Clearly, historical data are not useful. There is also no readily available analytical model (complicating factors include, for example, that the teams could not have played each other in their previous 12 games). Retroactive gambling line simulation is, however, easily applied. We conducted 100,000 simulation replications for each NHL season since 2005-6 (when shootouts were introduced to break ties) and recorded how many times this event occurred in each season. For additional insight, we recorded this for streaks lengths from 10 to 14 games. The results are shown in Table 6.

Table 6: Frequency of Games Between Two Teams With Long Win Streaks (100,000 Simulated Seasons)

	Win Streak Minimum (for both teams)				
Season	**10**	**11**	**12**	**13**	**14**
2005-6	266	90	31	12	5
2006-7	218	83	20	11	2
2007-8	92	26	5	3	0
2008-9	196	56	12	6	3
2009-10	153	41	12	5	0
2010-11	111	34	8	3	2
2011-12	132	31	14	3	0
2012-13	65	14	5	4	0
2013-14	167	51	18	3	0
2014-15	182	58	18	5	0
2015-16	133	24	16	3	1
Events Per Simulated Season	0.0015591	0.0004618	0.0001445	0.0000527	0.0000118
Seasons Per Event	641	2165	6918	18966	84615

The best estimate is likely the average across all seasons, which means that we would expect this event to happen about once every 6918 NHL seasons. That is pretty rare. We have a couple of other quick observations. First, the frequency is higher in seasons that had less competitive balance so this could actually serve as a measure of competitive balance (although there are much better ones). Second, we don't show results for streaks more than 14 games because this never occurred in 100,000 simulations of each of the 11 seasons. Such an occurrence would truly be exceptionally rare!

7 CONCLUSIONS

Retroactive gambling line simulation offers the capability to assess the probabilities of events that can't be assessed as accurately (or even at all) with analytical models. The ability to break the gambling lines into components and re-assemble them in various ways to best inform the desired analyses greatly extends the flexibility of the approach. Although the idea was motivated by the widespread discussions in the media comparing the 2015-16 Warriors and 1995-96 Bulls and we have used that as the expository application, it is the flexibility of the approach that we would like to emphasize. As is true with simulation in general, the number of what-if questions that can be assessed is large. The team performance, rare event perspective, and scheduling algorithm evaluation applications specifically identified here are just a start. As for the Warriors and the Bulls, the simulations show that, in addition to the obvious truth that 73 wins is more than 72, the Warriors performance was more impressive in the details as well.

REFERENCES

Barry, Daniel, and J.A. Hartigan, 1993. "Choice Models for Predicting Winners in Major League Baseball". *Journal of the American Statistical Association* 88 423:766-774.

Bowman, R. Alan, James Lambrinos and Thomas Ashman, . 2013, "Competitive Balance in the Eyes of the Sports Fan: Prospective Measures Using Point Spreads in the NFL and NBA". *Journal of Sports Economics* 14:498-520.

Bowman, R. Alan, Thomas Ashman, and James Lambrinos. 2013. "Prospective Measures of Competitive Balance: Application to Money Lines in Major League Baseball". *Applied Economics* 45:4071-4081.

Bowman, R. Alan, Thomas Ashman, and James Lambrinos. 2015. "Using Sports Wagering Markets to Evaluate and Compare Team Winning Streaks in Sports". *American Journal of Operations Research* 5: 357-366.

Cote, Jean, Dany J. Macdonald, Joseph Baker, and Bruce Abernethy. "When "Where" is More Important than "When:" Birthplace and Birthdate Effects on the Achievement of Sporting Expertise". *Journal of Sports Sciences* 24 10:1065-1073.

Glickman, Mark E., and Hal S. Stern. 1998. "A State-Space Model for National Football League Scores". *Journal of the American Statistical Association* 93 441:25-35.

Gupta, Ajay A. 2015. "A New Approach to Bracket Prediction in the NCAA Men's Basketball Tournament Based on a Dual-Proportion Likelihood". *Journal of Quantitative Analysis in Sports* 11 1:53-67.

Koop, Gary. 2002. "Comparing the Performance of Baseball Players: A Multiple-Output Approach, *Journal of the American Statistical Association* 97 459:710-721.

Lenten, Liam. 2016. "Mitigation of Perverse Incentives in Professional Sports Leagues with Reverse-Order Drafts". *Review of Industrial Organization* 49 1:25-41.

Levitt, Steven. 2004. "Why Are Gambling Markets Organized So Differently from Financial Markets?". *The Economic Journal* 114:223-46.

Paul, Rodney, and Andrew Weinbach. 2007. "Does Sportsbook.com Set Pointspreads to Maximize Profits? Tests of the Levitt Model of Sportsbook Behavior". *Journal of Prediction Markets* 1:209-218.

Paul, Rodney, and Andrew Weinbach. 2008. "Price Setting in the NBA Gambling Market: Tests of the Levitt Model of Sportsbook Behavior". *International Journal of Sport Finance* 4:225-263.

Rimler, Michael S., Seongho Song, and David T. Yi. 2010. "Estimating Production Efficiency in Men's NCAA College Basketball: A Bayesian Approach". *Journal of Sports Economics* 11 3:287-315.

Strumbelj, Erik, and Petar Vracar. 2012. "Simulating a Basketball Match with a Homogeneous Markov Model and Forecasting the Outcome". *International Journal of Forecasting* 28:532-542.

AUTHOR BIOGRAPHIES

R. ALAN BOWMAN is Professor of Operations and Information Systems and Senior Associate Dean of Graduate Business Programs at Clarkson University. He is located at the Clarkson University Capital Region Campus in Schenectady, NY. He holds a Ph.D. in Operations Research and Industrial Engineering from Cornell University. His research interests include scheduling, quality systems, project management networks, simulation, and sports economics. His e-mail address is abowman@clarkson.edu.

THOMAS ASHMAN is Associate Professor of Finance at Eckerd College in St. Petersburg, Florida. He teaches Finance and Strategic Management classes. His research is in sports economics and corporate finance. His email is ashmantd@eckerd.edu.

JAMES LAMBRINOS is Professor of Economics in the School of Business at Clarkson University's Capital Region Campus in Schenectady, New York. He has a Ph.D. in Economics from Rutgers University. His research interests are in health economics, forensic economics and sports economics. His email address is jlambrin@clarkson.edu.