Quantifying How NBA Playoff Format Impacts Playoff Excitement

Quinn Johnson

Summer 2020

Abstract

Throughout recent years, many professional American sports leagues have altered their postseason systems, but the National Basketball Association has predominantly kept its same basic format for decades: Each conference sends their top eight teams to the playoffs, a conference champion is crowned, and the two conference champions face-off in the best-of-seven NBA Finals. Here, we explore two alternative playoff formats and quantify the expected excitement in game play across the playoffs. We consider format changes to include each conference's top 8 teams and playing a conference-less tournament, as well as, taking the top 16 teams in the NBA and playing a conference-less tournament. After the simulations, we found small to no differences in excitement between the three formats but continue to endorse the top 16 team system due its advantage in fairness to all teams.

1 Introduction

In recent years, many professional American sports leagues have altered their postseason systems. Major League Baseball added the Wild Card game to their postseason in 2012 and the National Football League will be expanding their playoff field from 12 teams to 14 starting in the 2020-21 season. While the National Basketball Association (NBA) has made some minor changes over the last couple decades regarding divisions and seeding, the same basic playoff format has persisted.

In the National Basketball Association, 30 teams compete in an 82-game season from October to April in hopes of winning enough games to earn a postseason berth. There are two conferences, East and West, that both have 15 teams. Teams are ranked within their conference based on their overall wins throughout the season. If they finish in the top eight in their conference, they earn a spot in the NBA Playoffs. This implies that if the ninth seed in the West has 46 wins, they will not make the postseason, but if the eighth seed in the East has 45, they will. Situations like this are a relatively common occurrence in the league. In seven of the last ten postseasons (2010-2019), one conference has had a team eliminated with at least one more win than a team that earned a berth in the other conference.

Arguably a key goal of any playoff system from the perspective of the teams is fairness, which is discussed both in terms of which teams make the tournament and how they are seeded. For example, there are often discussions around whether a team actually deserves to be in the NBA Playoffs, especially if their win total is lower than a team in the other conference that was eliminated from postseason contention. The most recent time this scenario arose was in the 2017/18 season in which the Denver Nuggets sat in ninth place in the West after regular season play – and therefore didn't make the playoffs. However, Denver had a better record than three Eastern Conference playoff teams that earned the sixth (Miami Heat), seventh (Milwaukee Bucks), and eighth (Washington Wizards) seeds in the East. In addition, under the current format, a high seed in a strong conference is likely to have an unfairly difficult road to the NBA Finals since seeding is performed within conference.

Motivated by these concerns about fairness, two alternative playoff system proposals have been developed. The first sends the top 16 teams in the NBA into the postseason; we refer to this as the *16TOT system*. Under this format, in the 2017/18 postseason the Denver Nuggets would have been the 14th overall seed and the Washington Wizards would have missed the dance by one game. This format ensures that the league's best 16 teams compete in the playoffs, rather than the berths being equally split between conferences. While there would be disputes around the point of having conferences and divisions if the 16TOT system were employed, we argue this provides the best teams with the playoff berth they deserve, regardless of whether their conference opponents also have strong seasons. Therefore, the 16TOT format can be objectively viewed as the "best", most competitive format.

A second format would be that which includes the same teams as the current format, 8 from the East and 8 from the West, but seeds teams based on overall record. We refer to this as the *8W8E system*. Albeit a less ideal option because it does not resolve the conference strength dilemma, the 8W8E system does match-up the #1 overall seed with #16 overall seed, as opposed to maybe the #14 – which often happens under the current system. For example, if all eight teams in the Western Conference Playoffs are in the top 14 teams in NBA, the #1 seed in the West would play better teams than its counterpart in the East on the way to the NBA Finals.

Changing the system from its current form to a 16TOT format would drastically change the landscape of the playoffs. One way to quantify the degree of change incurred by altering formats is to count the difference between each playoff team's overall seed in the NBA, and their seed within the postseason. Aggregating these differences across all teams we obtain a sum of seed differences (SSD) metric which quantifies the impact of changing the system. For example, at the end of the 2011 NBA Season, the Indiana Pacers were the 8th seed in the East (16th seed in the playoffs), but were in 19th place in the overall NBA standings, corresponding to a seed difference of three. Additionally, the New York Knicks and Philadelphia 76ers were the 14th and 15th seeds, 15th and 16th overall, respectively. The 2011 NBA Playoffs had a SSD of five. In a perfect postseason situation, the top 16 teams overall will be made up of eight teams from each conference, thus having SSDs of 0 (e.g. 2012, 2017, 2019). Figure 1 shows the sum of seed differences for the last ten completed NBA postseasons (2010-2019). While some seasons wouldn't have different teams, the seasons with four or five summed seed differences would have playoffs that would look drastically different between the three formats. Figure 2 shows what the 2017 NBA Playoffs would look like under the three formats.

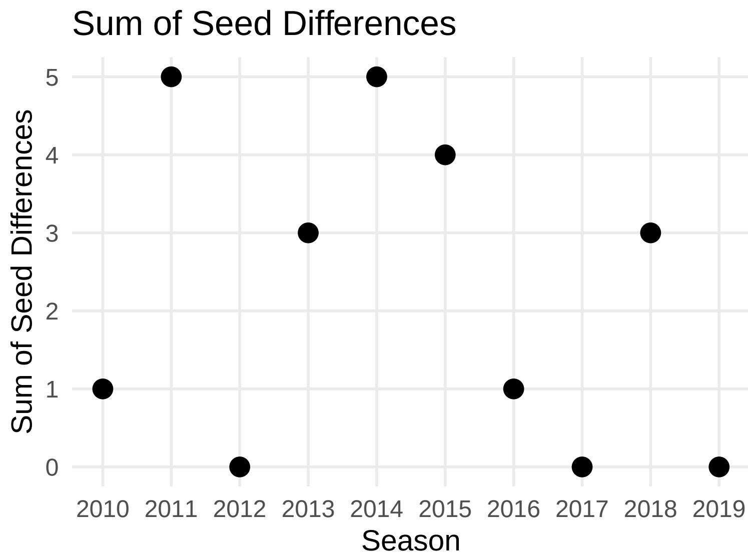


Figure 1: Sum of Seed Differences equals the sum of all playoff teams' overall seed in the NBA (1-30) and their seed in the playoffs (1-16). SSD > 0 indicates that 1 or more teams made the playoffs that should have been eliminated if only the top 16 teams in the league were given berths. The higher the SSD, the greater the differences in the playoff match-ups under the three formats.

While 8W8E gives berths to the same teams, playing a conference-less tournament greatly alters the match-ups and allows teams that could previously only meet in the Finals, to play in any round. The differences in match-ups between the current and 16TOT formats exemplifies truly how much altering seeding and adding the conference-less aspect transforms the postseason. While some may argue the 16TOT system is the more fair option, it is structured to give the best teams an easier road to the Finals, which could affect the excitement of the future NBA Playoffs.

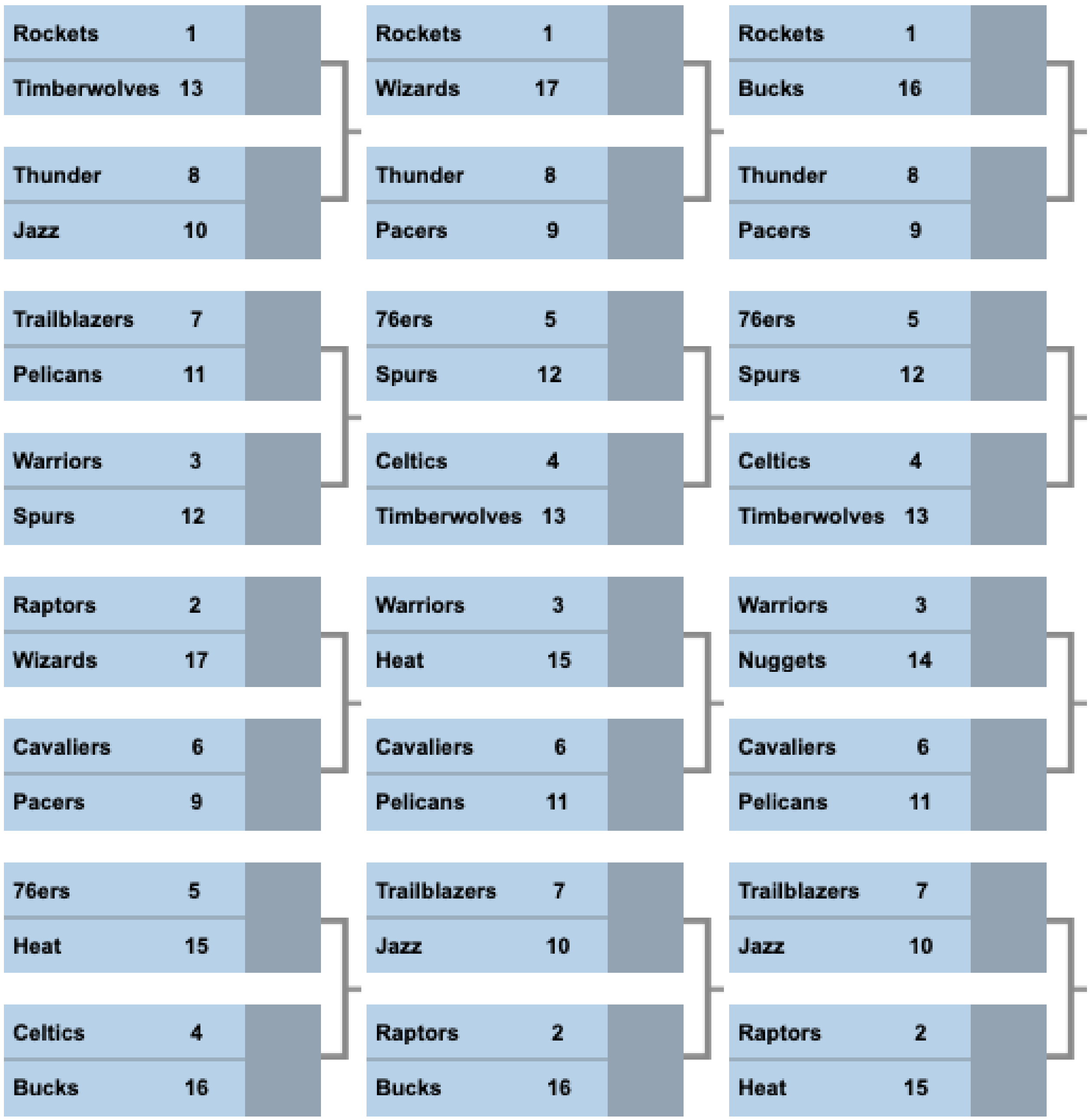
The NBA and its fans likely desire a playoff format that maximizes postseason excitement. It is good for the league if viewership is high and close games and series is one way to attract fans. Therefore, to compare the excitement value of the current playoff system to the 16TOT and 8W8E proposed systems, we used a Monte Carlo simulation to generate hypothetical playoff results under each scheme to determine if changing the format creates more exciting and competitive NBA Playoffs. This was quantitatively measured by simulating each format 10,000 times and tracking the number of games that (a) had a high likelihood of overtime/were expected to be close, (b) the lower seed in the game beat the higher seed, and (c) the number of series that went to six or seven games 1.

1.1 Data

The data used in this project are game scores from the 2016/17 through 2019/20 NBA seasons, scraped from Basketball-Reference.com [2] using the `RCurl`, `XLM`, and `tidyverse` packages in R [4]. Two large two-season-long data sets (2016/17-2017/18 and 2018/19-2019/20) were created containing every game that occurred during those four seasons. Each game log consists of eight variables with information about each game: date, start time, teams (home and away), points (home and away), overtime, and attendance. For the purposes of this analysis, the teams playing, location (home/away), and point difference were the primary focus (see Appendix for further details).

2 Methods

In 2018, Daniel Tokarz and the Yale Undergraduate Sports Analytics Group (YUSAG) developed their own NBA Power Rankings, which they used to predict win probabilities of hypothetical games [5]. Treating



(a) Current (b) 8W8E (c) 16TOT

Figure 2: 2018 NBA Playoffs First Round (Overall NBA Seeding). The match-ups immediately change once the tournament is made conference-less in the 8W8E format. Then, once the 16TOT format is applied, the Nuggets enter the picture which kicks out the Wizards and force the Heat and Bucks to play the 1 and 2 overall seeds, as opposed to the 2 and 3 seeds.

final score point differential of the teams' games as the response, they used a linear model (1) to estimate a value of each team's strength. After standardizing those coefficients to center around 0 and sorting them from highest-to-lowest, they obtained an objective, mathematical ranking of all 30 NBA teams. Next, the expected point difference of a fictional game was calculated with the estimated linear model using the strength coefficients of a team, their opponent, and whether the game was at home or not. Finally, Tokarz created a logistic model (2), that calculated the win probability of a game given the expected point differential. In the end, Tokarz/YUSAG's methods were able to rank the teams in the NBA and predict the outcome of a fictional, future game. Their work is the foundation of our models and simulations.

Here we discuss more in-depth the methodology (based off of YUSAG's work [5]) for modeling score differential and discuss how we use this model to simulate playoff game outcomes.

Expected point differential

To model a game's expected point differential, a linear model was fit using R's `lm()` function using ordinary least squares, where the point differential for game i (y_i) involving a primary team j and opponent k was modeled as a function of team's j's strength ($\beta_{T,j(i)}$), opponent k's strength ($\beta_{O,k(i)}$), and a home court advantage effect (β_H) based on whether team j was at home. This model can be expressed as

$$y_i = \beta_{T,j(i)} - \beta_{O,k(i)} + \beta_H \mathbb{1}_{\{Loc(i)=j\}} - \beta_H \mathbb{1}_{\{Loc(i)=k\}} + \epsilon_i \tag{1}$$

Each of the 30 NBA teams are represented by the subscripts j and k in the games they played for when they assume the "Team" perspective and when they are the "Opponent". As an example, if game $i = 1$ is between the Boston Celtics and Brooklyn Nets in Boston, then $j(1) =$ BOS so that the first term is $\beta_{T,BOS}$. Similarly, $k(1) =$ BKN so $\beta_{O,BKN}$ is the second term. Finally, $Loc(1) = BOS$ so the third term is β_H and the fourth term is zero.

Probability of winning

A logistic model was fit using the expected point differentials as the predictor, and the observed result of the game win or loss as the outcome. This model finds the a value, see (2), that minimizes the sum of squared error by fitting the logistic model that best fits the data. As a result, we obtain the expected win probability (W_p) for a team given the calculated expected point differential ($\hat{y}_i$)[5].

$$W_p = \frac{1}{1 + e^{-a \times \hat{y}_i}} \tag{2}$$

Using this probability of a win, the result of a new game can be simulated. To get expected point differential, the team, opponent, and location of a hypothetical game are used in the linear equation (1). Next, that point differential is plugged into the fitted logistic model (2) to find the expected win probability of this game. Finally, a random game outcome is generated based on this probability using a random Bernoulli sample with probability $p = W_p$. If the outcome of this Bernoulli sample is 1, the team won, if it's 0, the opponent won.

To simulate the 2018 postseason, games from the 2016/17 season and 2017/18 regular season were used to estimate the models. Then, to simulate the 2020 postseason, games from the 2018/19 season and 2019/20 regular season were used as training data. This two-season split was decided to be ideal due to the very volatile nature of small NBA team rosters and ensures that the games being simulated have as similar personnel as possible to the games used to train the model. Additionally, every team plays each other at least twice in a season, so that provides a minimum of four prior match-ups to use for the simulations.

Illustrating this with a fictional playoff game, suppose the Denver Nuggets host the Dallas Mavericks in Denver. The estimated linear model in (1) predicts the Nuggets to win this game by 0.86 points. This point differential corresponds to a 52% winning probability based on the estimated logistic model in (2). Thus, if the random Bernoulli sample with expected value of 0.52 draws a 0, the Mavericks end up getting the win at the Pepsi Center, and if the draw is 1, the Nuggets successfully defend their home court.

To simulate an entire postseason, this same process of gathering the teams and setting the location of the game was repeated thousands of times throughout the four rounds of the playoffs across the three formats. This generates a predictive distribution for the playoffs.

We compare the playoff formats by calculating each team's probability of winning the championship, and summarize overall playoff excitement using three metrics: upsets, number of overtime games, and number of long series (see Table 1). Playoff results for the 2018 postseason was simulated 10,000 times for each format. The 2018 postseason is used to determine the results rather than 2020 since all three formats would result in different brackets in this season (recall Figure 2), where only the current and 8W8E formats are applicable in 2020 due the top 16 teams in the NBA being equally split between the two conferences (8W8E is the same as 16TOT). The 2020 postseason was also simulated, and the results can be found in the GitHub Repository [3], along with 2018 results, code for scraping the data, simulating the games, and plotting the figures.

Metric	Description
Upset	Lower seed defeats higher seed
High Overtime	Game that is expected to be close ($\lvert .5 - W_p \rvert < .02$)
Long Series	Series that goes to Game 6 or 7

Table 1: Metric Descriptions

The tracked metrics were calculated separately for each by round and format and Tukey's Honestly Significant Difference (HSD) tests were conducted to determine which groups of metrics were statistically significantly different across formats.

3 Results

3.1 Championships

Before comparing the playoff formats based on the amount of excitement they generate using the simulated postseasons, we investigate the simulated postseason results to ensure they are reasonable. Figure 3 depicts the proportion of championships each team won of the simulated postseasons. In the 2018 NBA Playoffs, the Houston Rockets were the #1 overall seed, the Toronto Raptors were the #2, and the Golden State Warriors (the eventual NBA Champions) sat at #3.

There is a steep drop-off in championship probability after the Warriors and Raptors, which makes sense as the three top teams were the most dominant in regular season play. The Warriors played the Cleveland Cavaliers in the NBA Finals, so it is a little surprising to not see the Cavaliers have a larger share of simulated titles. Nonetheless, the overall trend of the simulations being a large difference between #3 and #4 match the Pre-Playoff Vegas Odds rather well (Table 2) as three teams were dominant and the rest had slim odds. The Warriors had very high pre-playoff odds due to them being the preseason favorites for the title after winning the championship in 2017 and finishing 73-9 in the 2015/16 campaign, but of course the simulation didn't take those things into account and was only given the information that they finished third in the NBA this season and first in 2017. Nonetheless, these results appear reasonable.

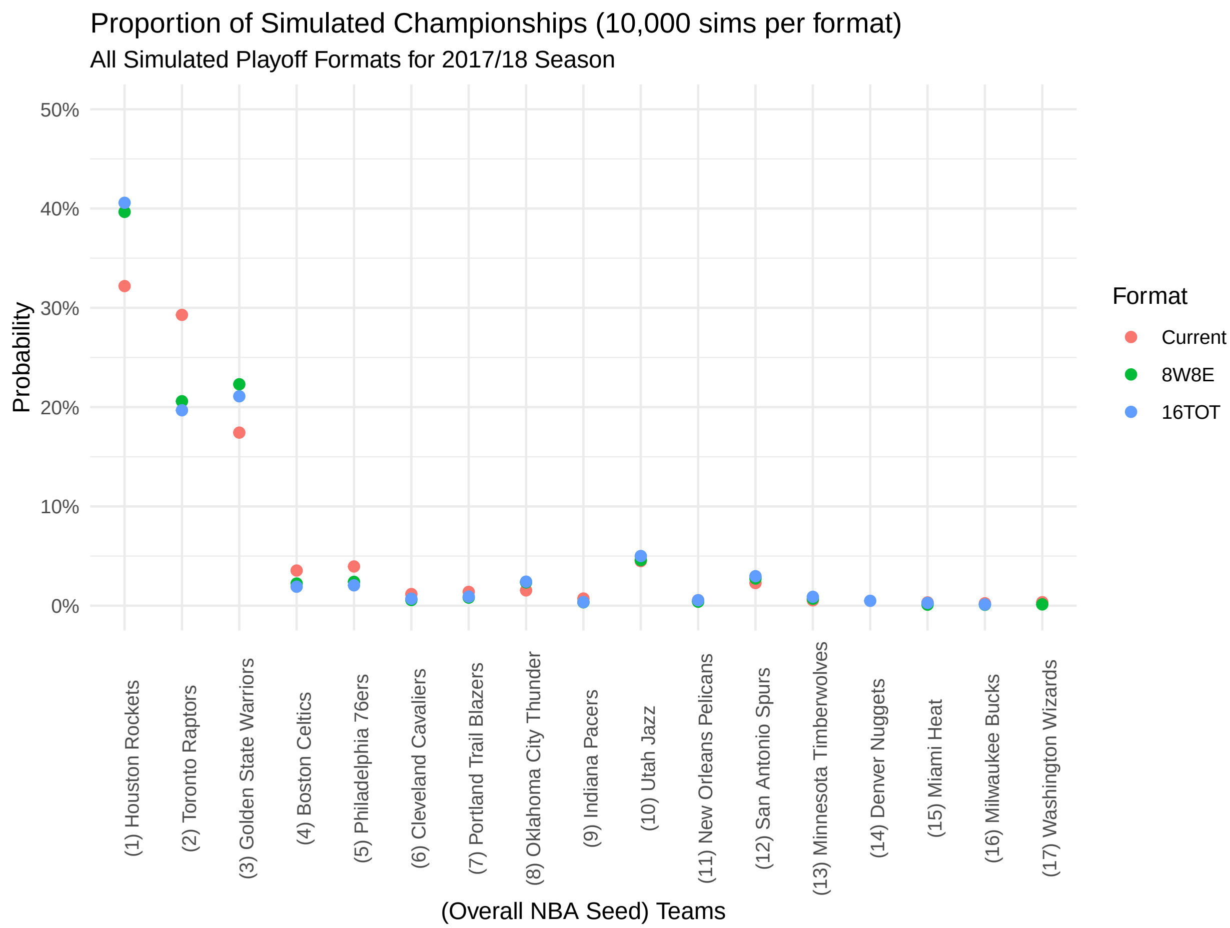


Figure 3: Probability of winning the 2018 NBA Championship under each playoff format. Clearly the Rockets, Raptors, and Warriors (actual champions) were the favorites. There are notable differences between the current results and the other two formats for the Rockets and the Raptors. Interestingly under the model, the Cavaliers were predicted to have a low probability of success even though they lost in the Finals. It is encouraging, however, that the Jazz have a significantly greater probability of winning than the higher-seeded Thunder, a team that the Jazz beat 4-2 in the first round.

Team	Estimated probability of Championship	Pre-Playoff Vegas Odds [1]
Rockets	0.32	0.42
Raptors	0.29	0.10
Warriors	0.17	0.44
Other	Less than 0.05	Less than 0.06

Table 2: Simulated vs Vegas Odds of winning the 2018 NBA Finals. The three favorites are the same, however, the Warriors are given a much better probability of winning by Vegas than by the simulations. This is due to the Warriors being dominant over prior seasons with proven playoff success, while the simulation is primarily using the fact that they finished 3rd overall (2nd in the conference) this season to make these predictions.

3.2 Upsets

The proportion of games that resulted in an upset are represented in Figure 4. Both the 2018 Eastern and Western Conference Finals finished four games to three, with the lower seeds, Warriors and Cavaliers, winning the series over the higher seeds, Rockets and Celtics, respectively. Therefore, there were eight upsets out of fourteen total games in this particular round (57%).

While a lot of the violins in Figure 4 extend far beyond practical values, there is still utility in comparing this, and the other excitement metrics, across formats.

Format	Mean Proportion of Upsets	HSD Group
Current	0.413	a
16TOT	0.407	b
8W8E	0.405	c

Table 3: Tukey's HSD Test for upsets. Each of these three formats are determined to be statistically different, however that is due to the very large sample size. They are in reality very close to each other with the difference being only about 1/2 more upsets per postseason between the current and 16TOT formats.

It is hard to tell a difference between some of the violins, and while the formats may be in different HSD groups (see Table 3 and Figure 7 in the Appendix), the means are, most of the time, shockingly close and only considered "significantly different" because of the extremely large sample size.

For upsets, it is fair to say that the current format produces more upsets than the other two, but by a slim margin. NBA Playoffs last on average around 83 games from start to finish. The 16TOT, 8W8E, and current format mean upset proportions are 40.7%, 40.5%, and 41.3%, respectively. Therefore, we can expect one more upset over a two-postseason span (1/2 more upsets per postseason) using the current format over the other two. This is rather negligible and not a clear argument for any of the three formats.

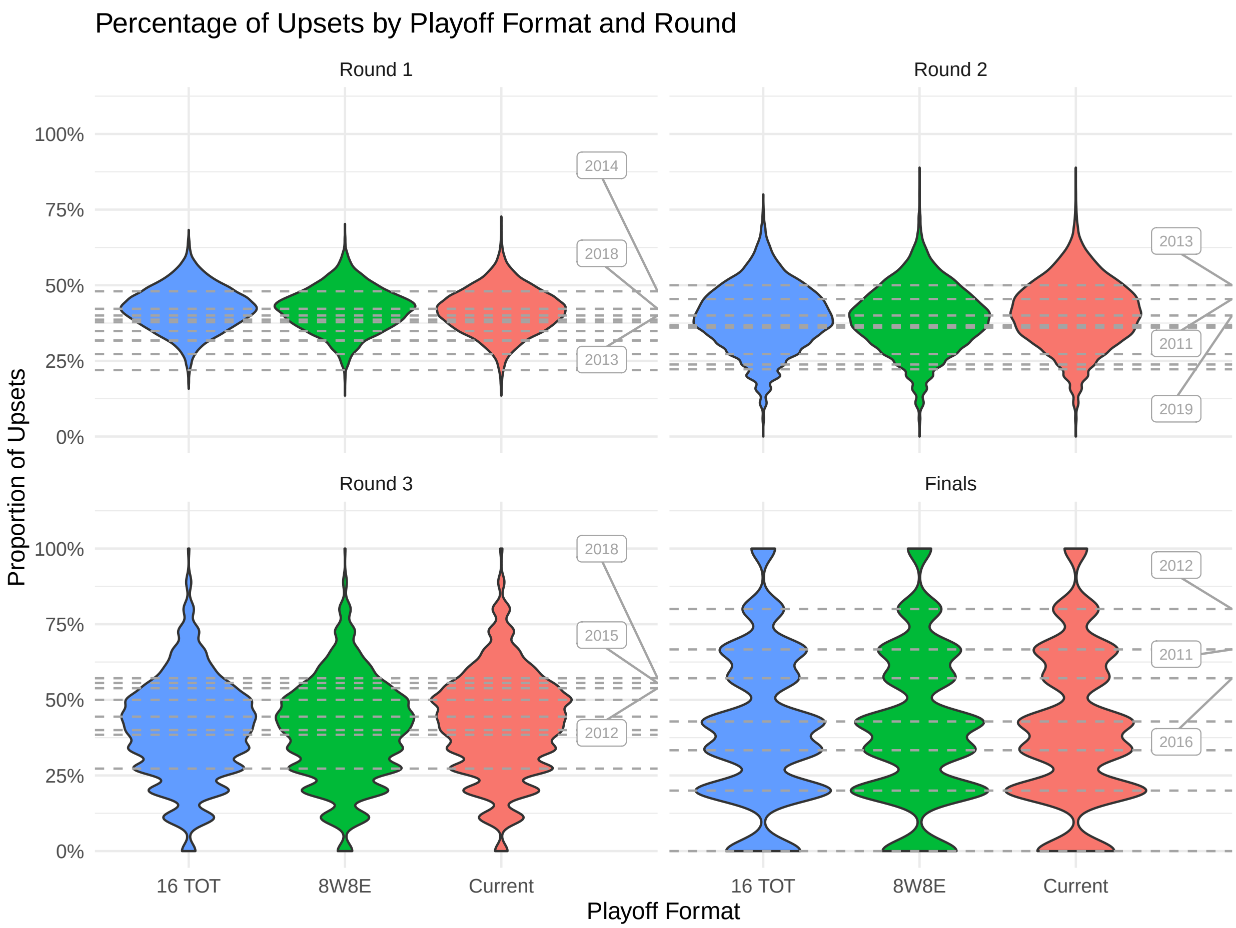


Figure 4: 2018 Proportion of simulated games ending in upsets by round and format. Many of the violins look almost identical; so much so, that the practical difference between the overall proportions by format only correlate to one more upset per two postseasons (1/2 more per postseason) under the current format. This difference is almost negligible considering that the hypothetical additional upset could come in any of the four rounds and be almost meaningless. Nonetheless, it is encouraging that none of the three formats produce significantly lower proportions of upsets.

3.3 High Overtime

Deriving the threshold for games that fall within the "high overtime" classification was difficult given the scores of the simulated games weren't reliable as actual results and were only used for finding a win probability. If the scores were realistic, the threshold would've been one possession (four points) to find close games, and tied at the end of regulation play for overtime games. However, it was decided that as long as the threshold was the same for all simulations and somewhat aligned with the proportion of actual NBA playoff games that go to overtime ($\sim 6\%$), then this metric would have some utility. It was thus decided that a game with an expected win probability of $.5 \pm .02$ (see Table 1) would be considered one with high probability of overtime.

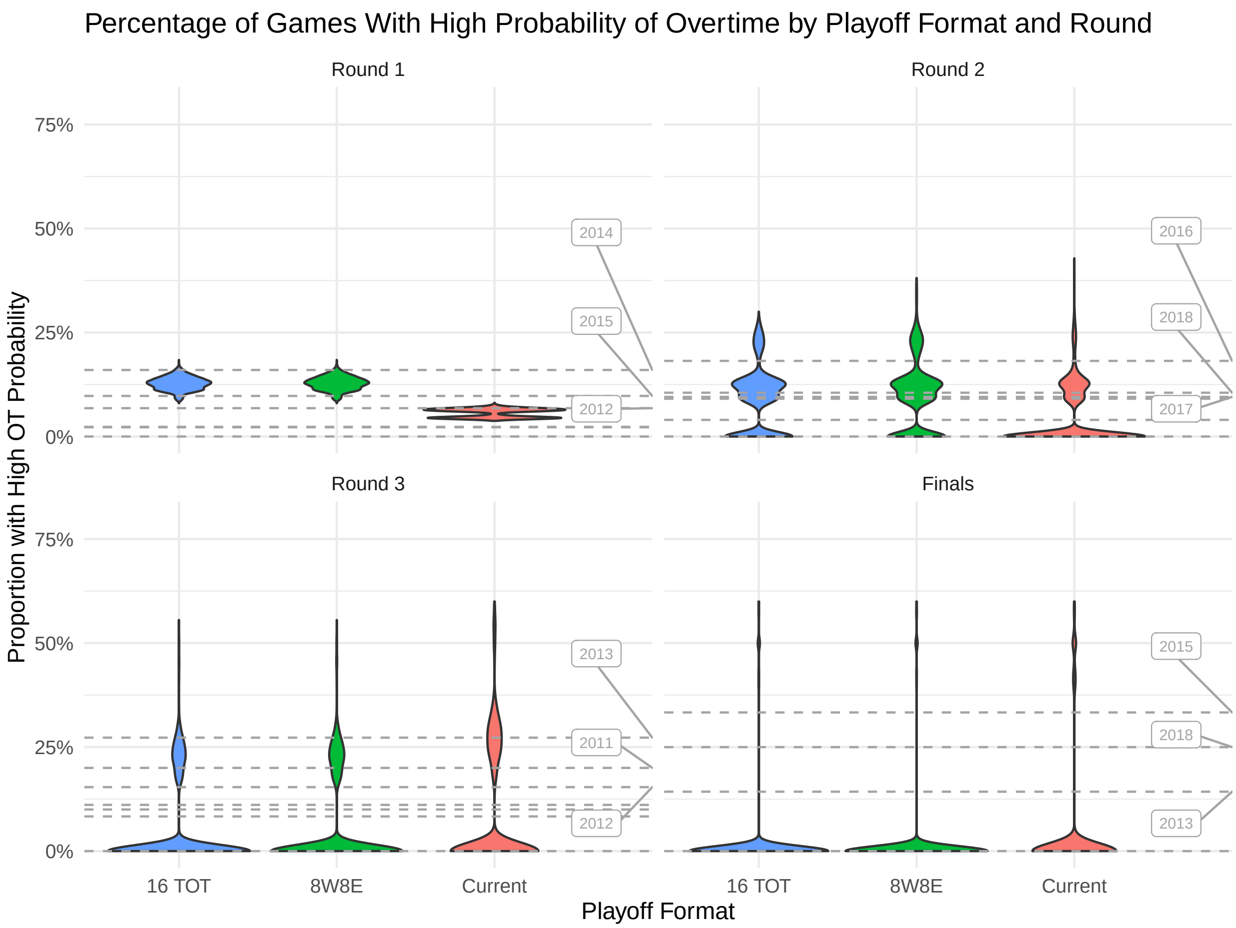


Figure 5: 2018 Proportion of simulated games with high probability of overtime by round and format. It is much clearer that there are greater differences in the violin plots for games with high likelihoods of overtime than it was with upsets. The greatest discrepancy is between the two conference-less formats and the current format in Round 1, with the current format proportions being visibly lower. As an NBA fan whose team may either be playing in one of the eight match-ups in the first round or not in the postseason at all, Round 1 can often be the least interesting and engaging as there are many other games going on and the stakes aren't as high as they are in the latter rounds. However, if more games had a higher probability of going to overtime/being close in crunch time, many more fans would be enticed to watch all the way through games. This is a clear result in favor of the conference-less formats.

Format	Mean Proportion of Games with High Probability of OT	HSD Group
8W8E	0.0755	a
16TOT	0.0718	b
Current	0.0598	c

Table 4: Tukey's HSD Test for games with high probability of OT. Like upsets, each of the three formats are determined to be statistically different, but this time the differences between the conference-less and the current formats is much more significant. The conference-less postseasons, according to the means, have on average 1.3 more close, potentially overtime, games per postseason than the current format. With each postseason only having a few overtime games per year, adding another can do nothing but improve the excitement and engagement of the playoffs, especially if that extra game is in a high stakes situation like a game seven or the Finals.

Figure 5 (and 8 in the Appendix) depicts the proportion of simulated games with a high probability of overtime/being close, with each round and format from the ten previous postseasons plotted for reference as well. Similarly to the upset results, the violins and outliers extend far beyond the realized overtime games, but this is to be expected from 10,000 runs of a simulation of this caliber. However, unlike upsets, the 8W8E and 16TOT violins are often as wider or greater than the current format high overtime probabilities. Therefore, the 8W8E and 16TOT format simulations produce more games that are closer/have a higher probability of going to overtime. The difference between the proportion means of 8W8E and current is about 1.6% (see Table 5), which results in about 1.3 more overtime games per postseason. This might not seem like much, but with some postseasons only have a couple overtime games over the whole tournament, adding one or more could be very exciting, especially if that one is a game seven or in the Finals.

3.4 Long Series

It is often said that "game seven" are the two most exciting words in sports. With every series in the NBA Playoffs being best-of-seven, every series has the opportunity to go to the sudden-death game. If a series is pushed to six games, then there is a good chance it could also go to seven as the teams are likely well-matched. Therefore, a postseason packed with series that extend to a game seven (or come close, at least), would be thrilling. Fans hope for close games and series that come down to the wire, and a game seven/sudden-death is the best way to decide those types of series. About 54% of NBA Playoff series currently go to six or seven games, but in the case of long series and the fan excitement they stir, it is beneficial to maximize the number of these thrilling games.

Format	Mean Proportion of Long Series	HSD Group
8W8E	0.5617	a
16TOT	0.5611	a
Current	0.5579	a

Table 5: Tukey's HSD Test for long series. Unlike the other two tests, all three formats were found to not be statistically significantly different, regardless of the large sample size. How truly similar the formats are can be seen in Figure 9, as they appear essentially identical. This is not an unimportant result, however, as it indicates that all three formats are capable of producing the same amount of long series, so there is no advantage to the current nor conference-less systems.

Figure 6 (and Figure 9 in the Appendix) display the simulated proportion of series that went to six or seven games, along with again, the previous ten completed postseasons for reference. It is almost impossible to determine which format has an advantage in the number of series that go long. However, it is important to note that no one format produces a significantly lower amount of long series. Therefore, we can reasonably expect all three formats to either hold the current standard of game sixes and sevens, if not exceed it.

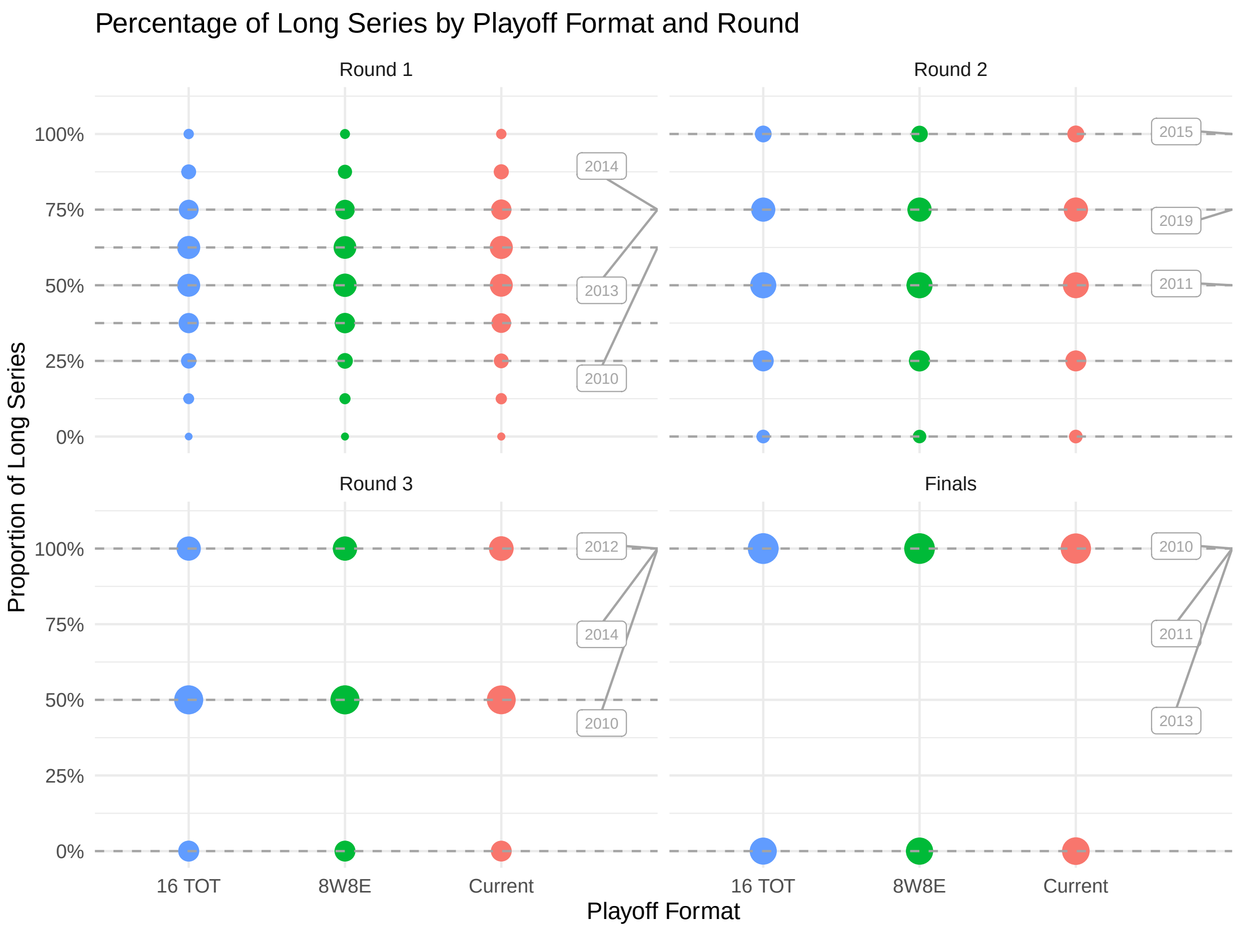


Figure 6: 2018 Proportion of simulated long series by round and format. Almost all of the proportions of series that went long in each round look identical across formats. However, it is encouraging that none of the formats under-perform in long series because this is arguably the most exciting metric.

4 Discussion

Whether the National Basketball Association should change their postseason format is an ongoing debate. There are arguments for and against each of the formats proposed in this exploration, as well as others that were not analyzed. To extend the conversation, other formats could be simulated, but it's impossible to know for certain what the impacts would be unless a change is actually made to the league's operations.

Next, this exploration was substantially limited by a few hurdles. Firstly, the expected point differences and win probabilities did not adjust and compensate for simulation results, meaning that Game 1 and Game 7 of a series would have the same prediction (only the random Bernoulli sample could differ). This is often far from the truth in real-life, as the match-up and team strategies evolve throughout a series. Another obstacle was processing power. It took hours to fully simulate each format after an adjustment, so altering code and quickly restarting full simulations was rarely a viable option. We were only able to reasonably perform a couple hundred simulations after each minor change to ensure the alteration was correct. This is due to the simulations following a step-by-step/Monte Carlo process that proved to be difficult to expedite. Lastly, it would be good to resolve the issue of failing to determine a way to accurately and consistently get realistic NBA game scores and point differences. It was difficult to model the randomness that is achieved by actual game results (the model predicted correct outcomes of the regular season games with 65.3% accuracy), but if a way to do so was found, there would be quite a bit more accuracy in game outcomes and more metrics could be analyzed with more depth. However, the NBA is notorious for its seemingly random results, and being able to accurately predict games' scores would be a lot more valuable in betting markets than simulating hypothetical postseason situations.

Finally, changing postseason formats is not only a discussion in the NBA. Almost every professional sport in America has an ongoing dialogue for how to improve the method in which a champion is crowned. However, for now, the NBA should adopt a format where the best 16 teams (or another reasonable number) in the league play in a conference-less tournament to name an NBA Champion. This is the most fair method as it puts every team on the same plane; no restrictions based on arbitrarily decided geographical boundaries. It plots the best teams in the league versus the ones that truly have to prove they deserve to be in the dance. From there the possibilities of Cinderella stories, nail-biting overtime games, and historical game sevens are endless and just waiting to be watched and enjoyed by NBA fans across the globe. Additionally, the 16TOT's competitiveness with – and occasionally outperforming of – the other formats in the three metrics prove that it would still be as or more exciting for the casual and die-hard fans, so it is wholly the ideal situation.

References

[1] NBA Futures Odds - Odds To Win The 2017-18 NBA Finals, 2018. https://web.archive.org/web/20180411202804/https://www.vegasinsider.com/nba/odds/futures/.

[2] Basketball Reference, 2020. https://www.basketball-reference.com/.

[3] Quinn Johnson. *NBAPlayoffExcitement GitHub Repository*, 2020. https://github.com/csu-sports-analytics/NBAPlayoffExcitement.

[4] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2020.

[5] Daniel Tokarz. NBA Model Math, 2018. https://sports.sites.yale.edu/nba-model-math.

A Data preparation

Before any analysis was conducted, the following variables were kept/created:

- **Date**: The date the game was played.
- **Location**: Whether the game was played at home ("H") or away ("A").
- **PtsDiff**: Difference between team points scored and opponent points scored.
- **Team**: The team in which the game's perspective is from.
  - If **Location** is "H", **Team** is the home team.
  - If **Location** is "A", **Team** is the away team.
    - Each game has two rows. One from the home team's perspective and one from the away team's. So **Team**, **Opponent**, **PtsDiff**, **Location** are opposite in the two rows.
- **Opponent**: Opposite of **Team**.
- **Win**: Binary indicator for whether the team won (1) or lost (0).
- **Days Since**: The number of days since the game occurred. The point of reference is determined by which postseason is being predicted and when that year's regular season ended. Used for weighting games in the linear model.
- **Season**: Differentiates in which season a game occurred. Used for weighting games in the linear model. This is an arbitrary constant that values the games within the season being simulated three times more than games from the season prior, so the values themselves could be changed if desired. However, it can be viewed as a way to increase sample size for the model (adding another season's worth of game results), while not giving those games the same influence that more recent games deserve. NBA team rosters, performance, and goals can vary greatly from season-to-season, so it is unreasonable to view games from over a year ago the same as the final game before the postseason begins. Because the NBA season begins in one year and ends in the next (i.e. the 2017/18 season began in October 2017 and ended in June 2018), the postseasons referred to in this report will be when the season ended/playoffs occurred (i.e. the 2017/18 playoffs will be "2018 playoffs").
  - If the game is in the season being simulated it should be weighted heavily, so **Season** $= 3$
  - If the game is in the season prior it should still be accounted for, but weighted less heavily and teams can change significantly from season-to-season, so **Season** $= 1$.
- **Game Weight**: Exponential function that weights games based on how recently they were played. Weights recent games more heavily, so they "count" for more in the model.
  - Season (1 or 3, depending on which one game i was played in) times the exponential of the inverse of the number of days since the game in question (game i) divided by the number of days since the game at the beginning of the data set (game 1)q.
  - $Season_i \times e^{-\frac{DaysSince_i}{DaysSince_1}}$

B Additional results

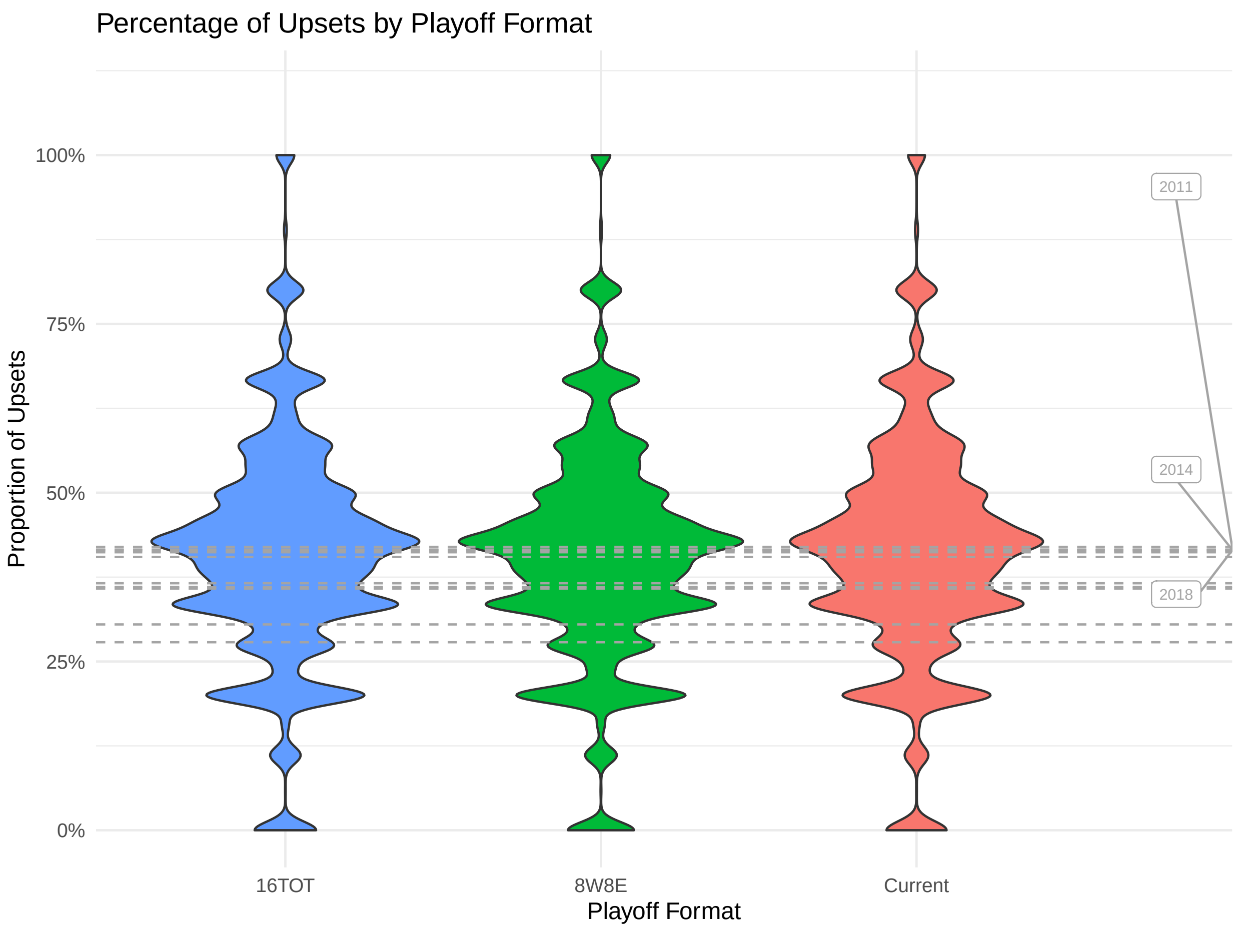


Figure 7: 2018 Proportion of simulate upsets by only format

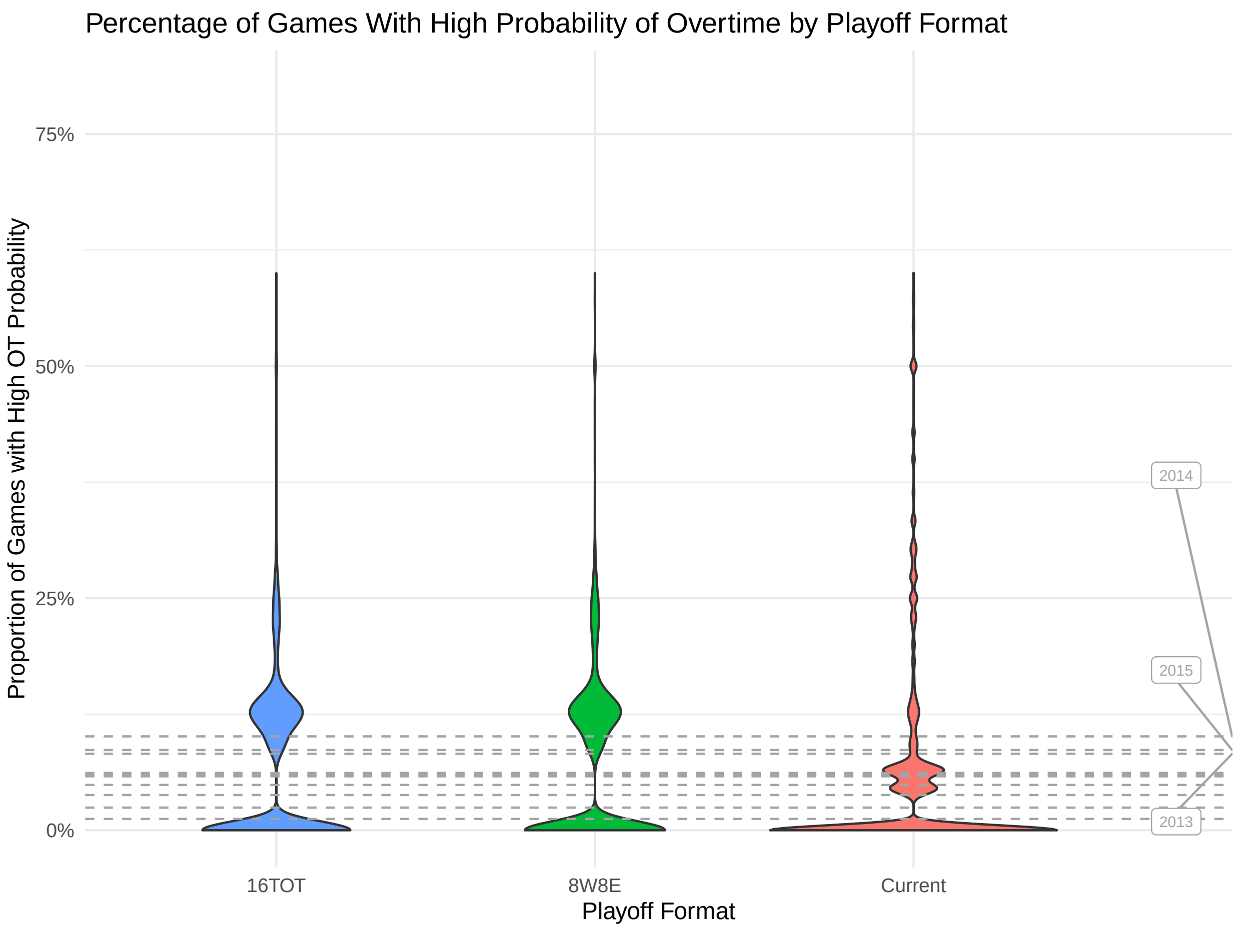


Figure 8: 2018 Proportion of simulated games with high probability of overtime by only format

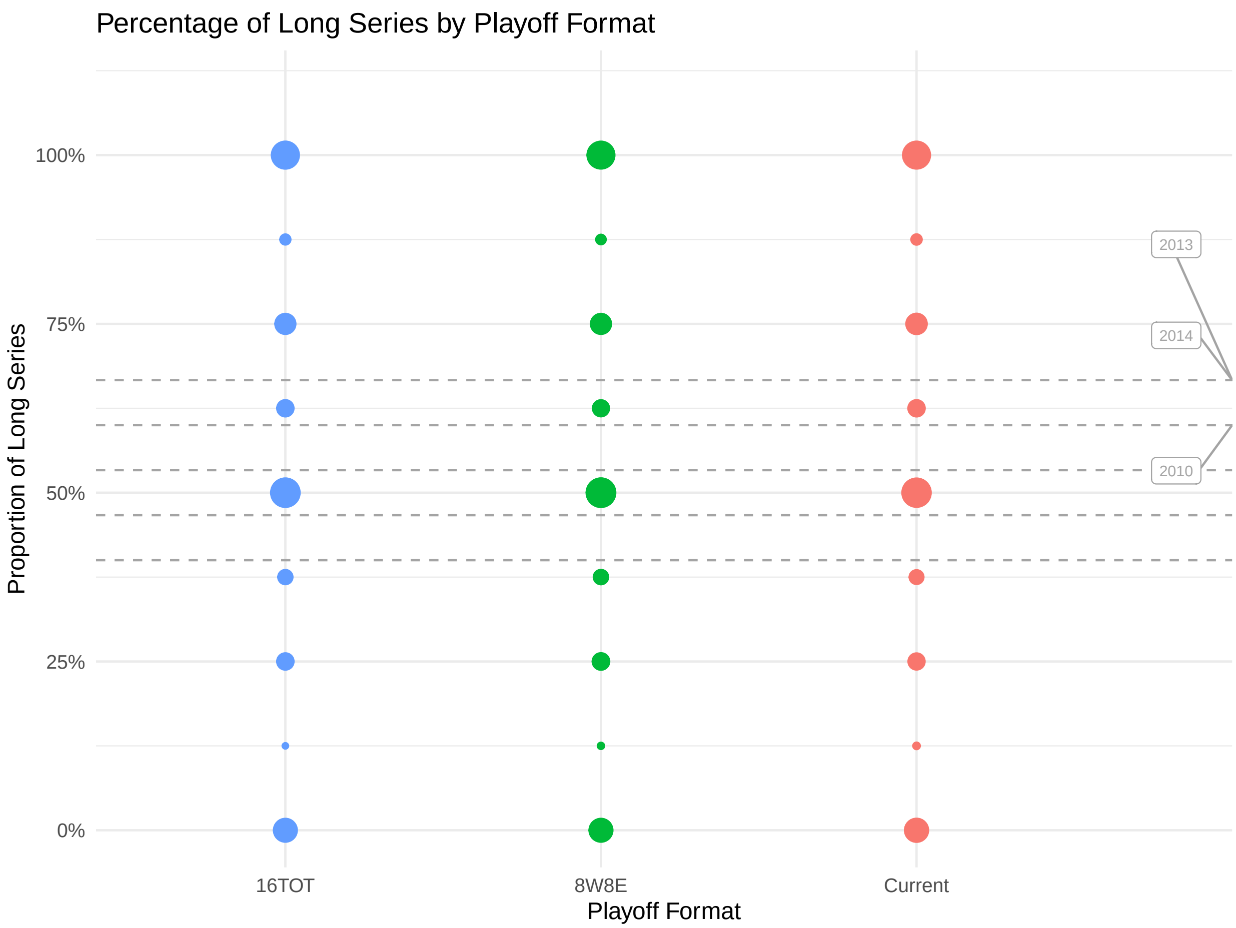


Figure 9: 2018 Proportion of simulated long series by only format