March Madness: Do Less Turnovers Lead to More Upsets?

May 28, 2024

Nathaniel Benner, Frank Pham, Jose Garcia, Miguel Serrano

Department of Statistics, Cal Poly SLO

Abstract

Between 2008 and 2024 (excluding 2020), upsets in the Men's March Madness tournament have increased with little explanation for why. One of the ideas presented as to why is turnovers, as teams that continue to get better and better start to turn over the ball less and less, leading to more opportunities to score and in turn giving the team a better chance of winning. In this research paper, we dove into the idea of turnovers having the biggest influence on the game of basketball and discovered convincing evidence that over the last 20 years, the decrease in average turnovers per game committed by a team forms a linear regression with the increase in the number of upsets in the March Madness tournament. By examining this correlation in detail, our analysis sheds light on how turnovers significantly impact game outcomes and contribute to the dynamic nature of March Madness, setting the stage for further exploration into other potential factors influencing these unexpected results.

1 Introduction

When considering the most attention-grabbing aspects of March Madness, a single-elimination tournament played in the United States to determine the men's college basketball national champion of the Division I level in the National Collegiate Athletic Association (NCAA), the one that comes to mind are the upsets every year. According to the NCAA, an upset is when the winning team in the NCAA tournament game is seeded at least five seeds worse than the losing team. An example is when number 16-ranked Fairleigh Dickinson triumphed over number 1-ranked Purdue in 2023, one of the biggest upsets in men's NCAA Basketball tournament history. Not only are upsets exhilarating to watch and one of the key attributes to predict when filling out a March Madness bracket, but they are also good stories to sell to the fans, as many people "end up voting for underdogs" (Fredén, 2020). Typically in the past, these upsets held more meaning, as they only occurred occasionally, aiding the fans' interest in them as they were more unlikely to happen. However, in the Men's March Madness tournaments post-2008, we have documented an unprecedented amount of upsets. One theory is due to a team's decrease in their turnovers on offense, as other researchers have found that "A decrease in turnovers by one percent increases the chances of an upset by 26%." (Ezekowitz, 2010). As described by the National Basketball Association, a turnover is when a team with possession of the basketball loses the ball to the other team before a shot is attempted. Turnover percentage (TOV%), similarly is the proportion of offensive possessions a team has in a game that results in turning the ball over to the other team out of every offensive possession a team has. Our goal in this paper is to determine whether the increase in upsets in March Madness post-2008 is correlated with a decrease in average turnovers across the men's college basketball landscape between 2008 and 2024.

2 Data and Methods

The following is a description of how our data was collected and the statistical analysis used on the data.

2.1 Data Description

Our data on March Madness from the years 2008-2024 (excluding 2020) was collected from numerous reputable sports companies, such as the Entertainment and Sports Programming Network (ESPN), American Broadcasting Company (ABC), and other data analysis websites. The subjects in our study were the men's college basketball teams that made it to the March Madness tournament and their team stats, gathered every year. These team stats included seed, field goals, field goal percentage, 3-pointers, 3-point percentage, free throws, free throw percentage, rebounds, offensive rebounds, defensive rebounds, assists, steals, blocks, turnovers, and fouls.

2.2 Statistical Analysis

We analyzed our data through a simple linear regression model where the variable 'Year' was modeled as a function of 'TOV%' (turnover percentage) within Rstudio and displayed the relationship between these two variables as well.

We also then graphed the percentage increase in upsets post-2008, meaning we took the number of upsets in 2008 (which was 9 out of a total of 63 games, the same number every year) and then compared the upset count for each year to 2008, noting the increase in the count by percentage. For example, in 2023 there were 14 upsets, which is about a 55% increase from the amount of 9 in 2008. We conducted an Exact binomial test to determine if the proportion of

games ending in an upset in 2023 is significantly different than the proportion in 2008, with n = 63 games taken through Rstudio. In R, we used the tidyverse, ggplot2, and readr packages to analyze the data in our tests below and create graphs that display the data effectively.

We will be comparing the number of upsets and how that number has changed over time compared to 2008, graphing the percentage increase in upset count between 2008 and 2009-2024. We will also be conducting a binomial test, questioning how likely the amount of 14 upsets (the upset count in 2023) would be, given the probability of an upset in 2008 (9/63, or about 0.143).

After comparing upset counts, we will then create a timeline of the average team turnover percentage across all teams in the tournament each year. A linear model will also be fit between Year and TOV% to see if the decrease in turnover percentage fits a negative linear model.

3 Results

With **Figure 1,** we can easily spot that post-2008, the number of upsets increased by at least 25%, except for 2009 and 2020 (no tournament). In **Table 1,** the exact binomial test provides weak evidence that the probability of an upset in 2023, 22%, is not the same as in 2008, as the probability of getting 14 successes or larger with the probability of success being 14% is about (p-value = 0.058).

In **Figure 2,** we can see the negative, linear relationship formed with Year and TOV% and **Table 2** proves that this relationship is significant, giving us convincing evidence of a decrease in TOV% between 2008 and 2023.

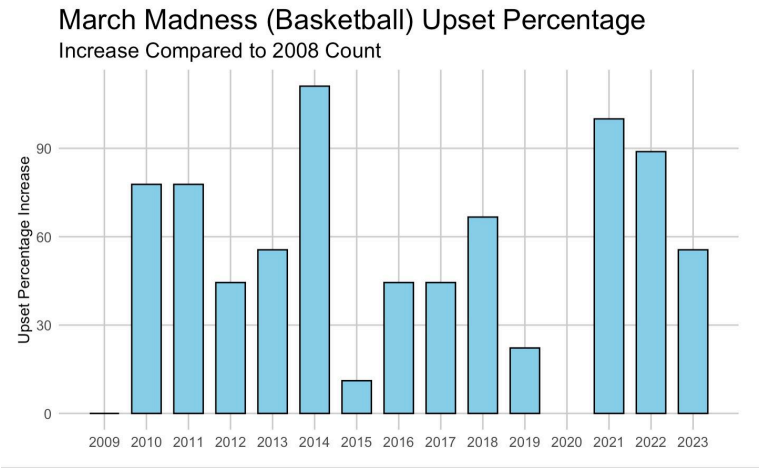

Figure 1: Upset Percentage Increase Between 2009-2023 Compared to 2008, showing that every year post-2008 saw an increase in the # of upsets compared to 2008

Successes	Trials	P-value	Conf. Interval	Sample P(Success)
14	63	0.05862	(0.140, 1)	0.222

Table 1: Binomial Test Results of the Probability of Getting 14 Upsets Given the 8 Upsets Observed in 2008

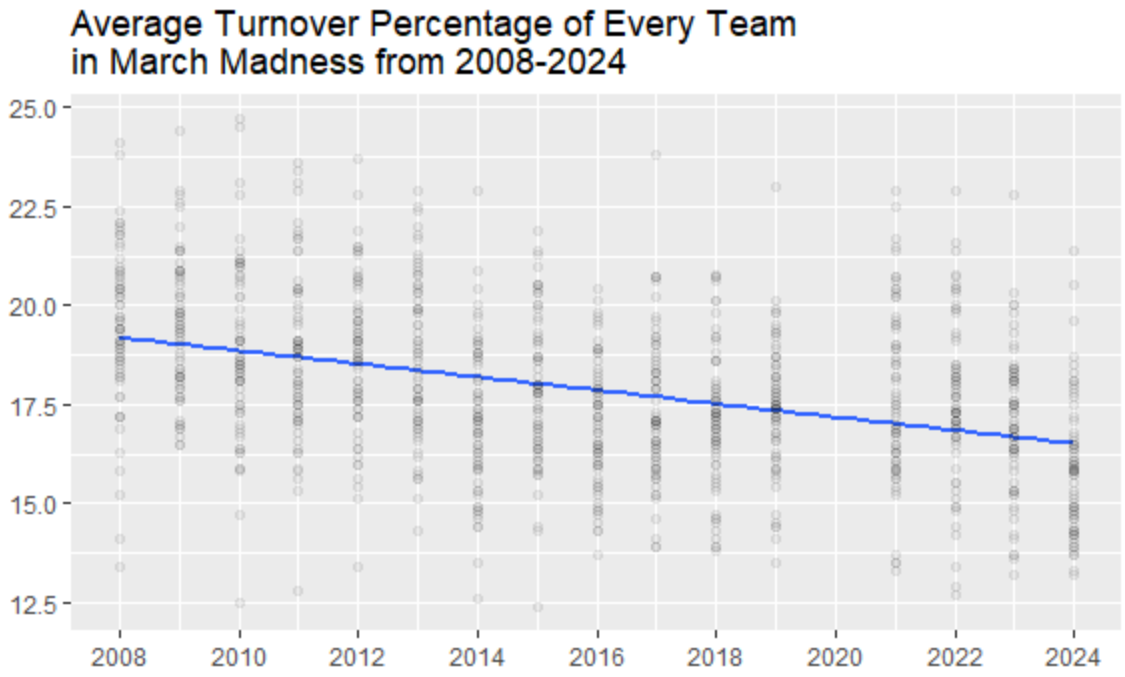

Figure 2: Linear Model of Turnover Percentage between 2008-2024

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 2031.97850    1.17491 1729.47   <2e-16 ***
`TOV%`        -0.90410    0.06523  -13.86   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 2: Linear regression test for Turnover percentage between 2008-2024

4 Discussion

Figure 1 shows the upset count percentage for the NCAA March Madness tournament from 2008 to 2023 over a total of 63 total played games. Where 2008 serves as the baseline that the following years are compared to. We observed how every year following 2008 there's been the same amount of upsets or an increase in the amount of upsets.

Furthermore, the average increase in upsets compared to 2008 across these 14 years was over 50%. Indicating that this increasing trend in upsets is consistent. This leads us to believe that the increasing trend in upsets is not just by chance and is a sign that other effects are at play including Average Turnover percentage which we will discuss its correlation with Upsets. The increase in Average Turnover percentage also signifies that the skill gap between teams has been tightening leading to the byproduct that is the increase in upsets.

According to **Table 1** above, we can observe how the probability of an upset in 2023 compared to an upset in 2008 has seemed to increase. However, we can not make any definitive conclusions due to our results being marginally insignificant ($x = 14$, $n = 63$, $p(\text{success}) = 0.143$, $p\text{-val}(\text{binom}) = 0.05862$) meaning we fail to reject the null that there is no evidence for the probability of an upset in 2023 being the same as in 2008.

Figure 2 displays a timeline of the average turnover percentage across all teams in the March Madness tournament. As we can see, the average turnover percentage in the tournament has been consistently decreasing, and the results of the simple linear regression test shown in **Table 2** support this, with the slope estimate being statistically significant (estimate = -0.904, t =

-13.86, df = 1077, $p < 0.001$). This proves that ever since 2008, a team's average turnover rate has decreased by about 0.9 of a percentage point each year.

This steady decline in turnover percentage may be a key factor contributing to the increased frequency of upsets observed in recent March Madness tournaments, as the less a team turns the ball over, the more opportunities they get to score, leading to smaller teams to have more chances to upset the heavily favored powerhouse universities.

5 Conclusion

After looking at the data following the upset counts, we can notice a positive relationship when it comes to the passage of time and the increased amount of upsets in a single tournament, coupled with the data evidence of a significant decrease in turnover percentage over time between 2008 and 2024.

Grounding this in the world of basketball, turnovers have already been discussed as one of the main contributing factors when it comes to winning basketball games. In a research study conducted investigating the main contributors when it comes to winning, researchers discovered that "When examined on a professional level of European basketball competition, winning teams were able to attempt and make more free-throw, 2-point, and 3-point shots as well as have fewer turnovers" (Cabarkapa, 2022). This supports that turnovers are well known to be correlated with success in the world of basketball, as they are "important because a team cannot even shoot, never mind score, if it turns over the ball during a possession." (Strauss, 2019).

In the future, we hope teams use this data from their past to better tailor their game towards the most optimal way to win, as that is all their fans want them to do.

References

Ezekowitz, J. (2013, January 12). Predicting NCAA tournament upsets: The underdogs. The Harvard Sports Analysis Collective. https://harvardsportsanalysis.wordpress.com/2010/03/15/predicting-ncaa-tournament-upsets-the-underdogs/

Fredén, A., Rheault, L., & Indridason, I. H. (2020). Betting on the underdog: the influence of social networks on vote choice. Political Science Research and Methods, 10(1), 198–205. https://doi.org/10.1017/psrm.2020.21

Cabarkapa, D., Deane, M. A., Fry, A. C., Jones, G. T., Cabarkapa, D. V., Philipp, N. M., & Yu, D. (2022). Game statistics that discriminate winning and losing at the NBA level of basketball competition. PloS One, 17(8), e0273427. https://doi.org/10.1371/journal.pone.0273427

Strauss, F. L. &. L. (n.d.). What Wins Basketball Games, a Review of "Basketball on Paper: Rules and Tools for Performance Analysis" by Dean Oliver. Strauss, Factor, Laing & Lyons. https://www.sfandllaw.com/articles/what-wins-basketball-games-a-review-of-basketball-on-paper-rules-and-tools-for-performance-analysis/