title

Stanley Zheng Eric Lucas *

April 27, 2025

1 Abstract

This project investigates how the relationship between three-point shooting and team success in the NBA has evolved over the past two decades. Using team-level data from the 2003-04 and 2023-24 regular seasons obtained via the NBA API, we construct multiple linear regression models to examine the effects of three-point percentage (3P%), two-point percentage (2P%), and shot selection proportions on win rate. Our analysis reveals a statistically significant interaction between 3P% and era, indicating that the marginal impact of three-point accuracy on team success has increased substantially in the modern NBA. While 3P% showed no significant association with win rate in 2003-04, it has become a strong predictor in 2023-24—even after controlling for other efficiency metrics. This shift reflects a broader structural change in league strategy, marking the emergence of a "three-point era." Though based on observational data, the use of interaction terms allows for a quasi-causal interpretation of this evolving relationship.

2 Introduction

2.1 Motivation

Over the past two decades, the NBA has undergone a dramatic transformation in offensive strategy. The rise of "small-ball" systems and analytics-driven decision making has led to a surge in three-point attempts, reshaping how teams space the floor, select shots, and build rosters. Players like Stephen Curry have redefined the value of long-range shooting, prompting coaches and front offices to reconsider the role of the three-point shot in winning games. This project is motivated by a key question at the heart of this shift: Has the importance of three-point shooting truly increased over time, and if so, how does it compare to traditional metrics like two-point efficiency or shot selection

*Equal contribution

proportions? By using regression analysis on team-level data from different seansons, we aim to quantify how the statistical relationship between shooting performance and win rate has changed. In the era of big data, we believe that understanding these evolving dynamics can help teams identify actionable areas for improvement and optimize offensive strategies for greater success.

2.2 Dataset

Our dataset is sourced from the NBA's public API, an online interface that provides comprehensive and standardized statistics for teams and players across multiple seasons. The NBA API offers detailed data covering a wide range of performance categories, including shooting statistics, rebounding, passing, turnovers, fouls, player efficiency metrics, and advanced team analytics.

The full dataset records team-level aggregates for each regular season, capturing key indicators such as field goal percentages, three-point and two-point shooting volume and accuracy, free throw statistics, rebound counts, assist counts, turnover rates, steal and block numbers, and overall team performance metrics like win-loss records and plus-minus ratings. The structure of the dataset is shown in the Figure 1, and the detailed explanation of each variable is shown in the Appendix.

	TEAM_ID	TEAM_NAME	GP	W	L	W_PCT	MIN	FGM	FGA	FG_PCT	FG3M	FG3A	FG3_PCT	FTM	FTA	FT_PCT	\
0	1610612737	Atlanta Hawks	82	28	54	0.341	3981.0	2829	6529	0.433	419	1249	0.335	1534	1976	0.776	
1	1610612738	Boston Celtics	82	36	46	0.439	3941.0	2843	6415	0.443	553	1599	0.346	1572	2095	0.750	
2	1610612741	Chicago Bulls	82	23	59	0.280	3966.0	2798	6753	0.414	429	1256	0.342	1330	1834	0.725	
3	1610612739	Cleveland Cavaliers	82	35	47	0.427	3971.0	2922	6753	0.433	247	786	0.314	1528	2030	0.753	
4	1610612742	Dallas Mavericks	82	52	30	0.634	3961.0	3322	7230	0.459	507	1456	0.348	1475	1854	0.796	

	OREB	DREB	REB	AST	TOV	STL	BLK	BLKA	PF	PFD	PTS	PLUS_MINUS	GP_RANK	W_RANK	L_RANK	W_PCT_RANK	MIN_RANK	\
0	996	2507	3503	1648	1350.0	627	408	416	1826	3	7611	−381.0	1	25	25	25	4	
1	851	2440	3291	1683	1332.0	770	331	408	1837	2	7811	−119.0	1	20	20	20	28	
2	1048	2519	3567	1793	1321.0	659	396	446	1911	3	7355	−521.0	1	28	28	28	11	
3	1118	2619	3737	1808	1216.0	585	537	434	1743	2	7619	−215.0	1	21	21	21	7	
4	1174	2538	3712	1963	1000.0	656	437	382	1609	3	8626	364.0	1	7	7	7	15	

	FGM_RANK	FGA_RANK	FG_PCT_RANK	FG3M_RANK	FG3A_RANK	FG3_PCT_RANK	FTM_RANK	FTA_RANK	FT_PCT_RANK	OREB_RANK	DREB_RANK	\
0	17	16	20	15	14	23	12	17	6	16	10	
1	15	22	10	3	3	12	8	8	18	27	17	
2	20	4	29	12	13	19	27	23	25	8	9	
3	10	4	21	29	28	29	13	12	14	3	4	
4	1	1	3	6	5	11	15	22	1	1	6	

	REB_RANK	AST_RANK	TOV_RANK	STL_RANK	BLK_RANK	BLKA_RANK	PF_RANK	PFD_RANK	PTS_RANK	PLUS_MINUS_RANK
0	12	22	26	17	12	15	23	15	15	26
1	27	17	24	2	25	12	24	24	8	19
2	4	10	23	13	15	22	28	15	26	28
3	1	8	15	23	3	19	15	24	14	22
4	2	3	1	15	10	7	3	15	1	6

Figure 1: The Description of The Original Dataset.

For our analysis, we primarily focus on shooting-related metrics, including three-point field goal percentage (3P%), two-point field goal percentage (2P%), three-point attempt rate (3PA rate), and two-point attempt rate (2PA rate). Three-point and two-point attempt rates are computed as the proportion of total field goal attempts originating from each respective shot zone. To control for other aspects of team strength that may confound the relationship between shooting performance and win rate, we also incorporate variables such as total rebounds, turnovers, and other team-level performance indicators as covariates in our regression models. This approach allows us to better isolate the specific impact of shooting metrics on team success across different eras.

3 Data Analysis

3.1 Data Preprocess

Our data preprocessing involved three major steps. First, we retrieved the raw data from an API, where each table corresponds to a different season, and each row records a team's performance for that year. Therefore, our unit of analysis is team/year.

As shown in Figure 1, a large portion of the raw variables consisted of rankings (e.g., team rank in various metrics), which we decided to remove, as they are not suitable for our analysis. We also observed that the raw data already included some basic performance statistics, such as win percentage (W_PCT) and three-point shooting percentage(FG3_PCT). However, since the available statistics were limited to only a few dimensions, we manually calculated additional variables: win rate(WinRate), three-point attempt rate(3P_ratio), three-point shooting percentage(3P%), and two-point shooting percentage(2P%).

We found that the three-point attempt rate and the two-point attempt rate are in a perfect linear relationship (they add up to 1), so we included only the three-point attempt rate in our models to avoid multicollinearity.

Finally, because most of our key variables, such as win rate and shooting percentages, naturally fall within the range [0, 1], we applied normalization to the other numerical variables as well to ensure these variables also fall with in the range [0, 1], which maintains the consistency across features.

3.2 Data Visulization

To motivate our investigation into the evolving impact of three-point shooting on team success, we first examine two sets of histograms comparing the 2005-2009 and 2020-2024 NBA regular seasons.

The first set of histograms(Figure 2) shows the three-point shooting percentage (3P%), capturing efficiency on three-point attempts. While the distributions for the two periods overlap significantly, there is a modest rightward shift, indicating that teams have become slightly more accurate from beyond the arc over time. This suggests that the increase in three-point volume has not come at the expense of efficiency, but rather, teams have adapted by developing better shooters and improving shot quality.

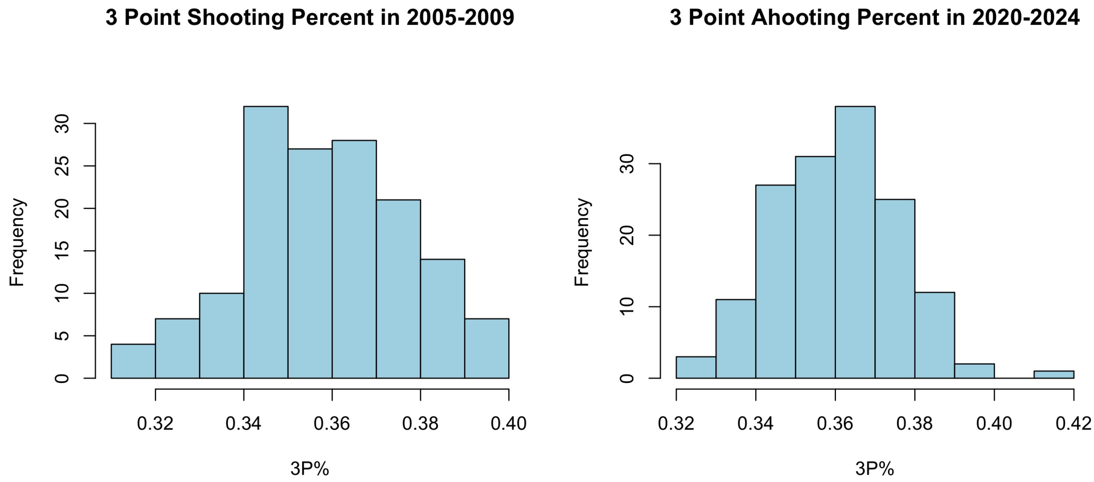

Figure 2: 3 Point Shooting Percent in Two Different Period.

The second set of histograms(Figure 3) displays the three-point attempt rate (3P_ratio), defined as the proportion of a team's total field goal attempts that are three-pointers. We observe a substantial rightward shift over time: in the 2005–2009 period, teams clustered around a 20% attempt rate, whereas by 2020-2024, the distribution centers around 38-40%. This sharp increase visually confirms the commonly cited notion that the NBA has entered a "small ball" era, characterized by greater reliance on perimeter shooting and faster-paced offenses.

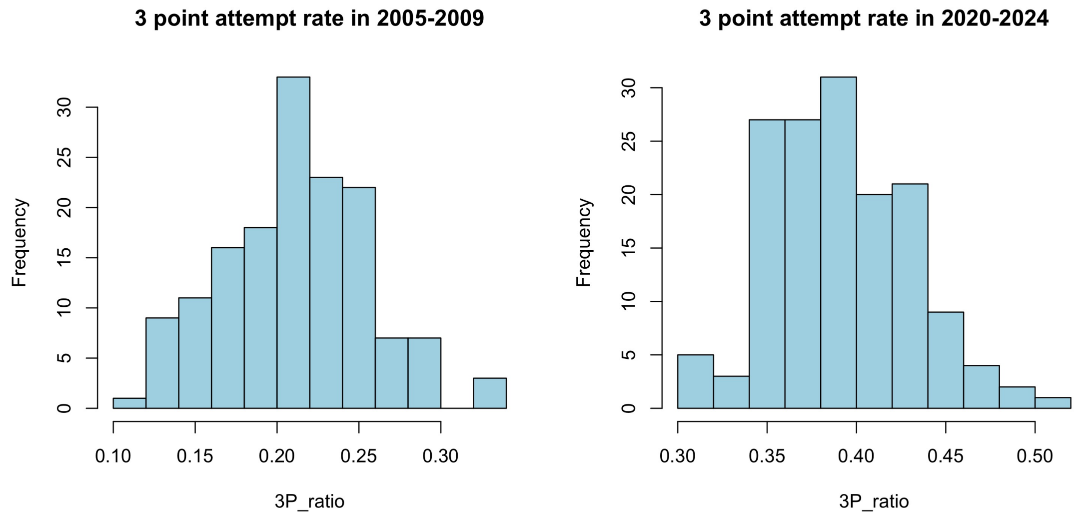

Figure 3: 3 Point Attempt Rate in Two Different Period.

Together, these visualizations provide structural evidence for a fundamental change in league-wide offensive strategy. The sharp increase in three-point attempt rates, coupled with stable or slightly improved three-point accuracy, sets the stage for our central analysis: evaluating whether and how the growing emphasis on three-point shooting translates into higher win rates in the modern NBA.

4 Model

4.1 Model Specification

To examine how the relationship between shooting efficiency and team success has evolved over time, we first fit a parsimonious linear regression model in each era, using team win rate (WinRate) as the dependent variable. The explanatory variables included three-point field goal percentage (3P%), two-point field goal percentage (2P%), and three-point attempt rate (3P_ratio).

$$\text{WinRate} = \beta_0 + \beta_1 \times \text{3P\%} + \beta_2 \times \text{2P\%} + \beta_3 \times \text{3P\_ratio} + \epsilon$$

This minimal specification allows us to isolate the marginal effects of shooting efficiency without the confounding influence of other team characteristics.

We separately fit the model to two distinct periods: the 2005–2009 seasons and the 2020–2024 seasons. The regression results reveal clear structural differences between eras.

In the 2005–2009 model, two-point field goal percentage (2P%) is highly significant ($p \approx 3.11 \times 10^{-08}$) and exhibits a strong positive relationship with win rate (coefficient = 3.7431). Three-point field goal percentage (3P%) is statistically significant at the 1% level ($p \approx 0.0038$), which is just a 2 stars significant in R, and its impact (coefficient = 1.8842) is noticeably smaller than that of two-point efficiency. Three-point attempt rate (3P_ratio) shows no statistical significance ($p \approx 0.301$). The detailed model summary is shown in Figire 4 left.

```
Call:
lm(formula = WinRate ~ X3P. + X2P. + X3P_ratio, data = mydf_2005_2009)

Residuals:
      Min        1Q    Median        3Q       Max
-0.286612 -0.091560  0.009916  0.080742  0.261863

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.0312     0.3044  -6.672 4.87e-10 ***
X3P.          1.8842     0.6407   2.941  0.00381 **
X2P.          3.7431     0.6399   5.850 3.11e-08 ***
X3P_ratio     0.2650     0.2554   1.038  0.30118
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1216 on 146 degrees of freedom
Multiple R-squared:  0.3694,	Adjusted R-squared:  0.3564
F-statistic: 28.51 on 3 and 146 DF,  p-value: 1.436e-14
```

```
Call:
lm(formula = WinRate ~ X3P. + X2P. + X3P_ratio, data = mydf_2020_2024)

Residuals:
      Min        1Q    Median        3Q       Max
-0.282708 -0.060843 -0.004748  0.068760  0.251019

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.3217     0.2431  -9.551  < 2e-16 ***
X3P.          4.7683     0.5728   8.324 5.58e-14 ***
X2P.          2.1837     0.4371   4.996 1.65e-06 ***
X3P_ratio    -0.1782     0.2271  -0.784    0.434
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1022 on 146 degrees of freedom
Multiple R-squared:  0.4985,	Adjusted R-squared:  0.4882
F-statistic: 48.38 on 3 and 146 DF,  p-value: < 2.2e-16
```

Figure 4: model0 2005 2009.

By contrast, in the 2020–2024 model, three-point field goal percentage (3P%) becomes even more strongly significant ($p \approx 5.58 \times 10^{-14}$) and its marginal effect size increases substantially (coefficient = 4.7683). Meanwhile, the coefficient on two-point field goal percentage (2P%) decreases to 2.1837, although it remains statistically significant. Again, three-point attempt rate (3P_ratio) remains statistically insignificant. The detailed model summary is shown in Figire 4 right.

These patterns suggest a fundamental shift:

- In the earlier period, improving two-point shooting efficiency had a larger marginal benefit to team success than improving three-point efficiency.

- In the modern period, improving three-point shooting efficiency has become more important than two-point efficiency, with its estimated impact on win rate more than doubling the impact of two-point shooting.

This transition supports the hypothesis that the NBA has entered a "three-point era," where teams' success increasingly depends on their ability to make three-point shots efficiently rather than solely relying on two-point scoring.

More than these two models, I also fit another two models in the periods 2010-2014 and 2015-2019. Using the data from these four models, I have create a table and a line chart. You can see the strong trend that 3P% is more and more important with the time goes in the Table 1 and the Figure 5.

Variable	2005-2009	2010-2014	2015-2019	2020-2024
3P%(coefficient)	1.8842	2.6011	3.8326	4.7683
2P%(coefficient)	3.7431	4.0901	3.1201	2.1837
3P%(significance)	0.00381	4.09×10^{-6}	1.13×10^{-9}	5.58×10^{-14}
2P%(significance)	3.11×10^{-8}	1.64×10^{-12}	1.51×10^{-8}	1.65×10^{-6}
3P_ratio(significance)	0.30118 (n.s.)	0.191 (n.s.)	0.0966 (n.s.)	0.434 (n.s.)

Table 1: Comparison of Marginal Effects Between Periods

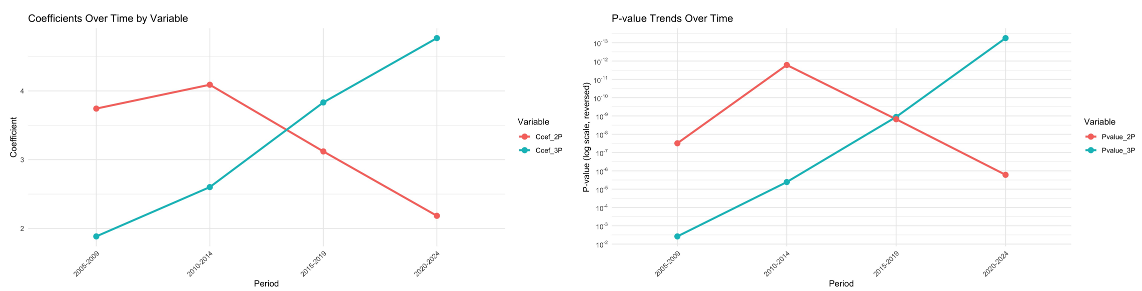

Figure 5: Comparison of Marginal Effects Between Periods.

4.2 Model Diagnostic

For our parsimonious linear regression models, which is actually our first type of model. We conducted standard model diagnostic checks using residual plots. The Residuals vs Fitted plot shows no strong patterns, supporting the linearity assumption. The Q-Q plot indicates that residuals are approximately normally distributed. The Scale-Location plot suggests constant variance of residuals. Finally, the Residuals vs Leverage plot does not show any extreme influential observations. Overall, the diagnostics suggest that our model is appropriate for inference. You can see the diagnostic plots in Figure 6, this is the polt for the period 2005-2009. For different periods, the diagnostic plots are similar.

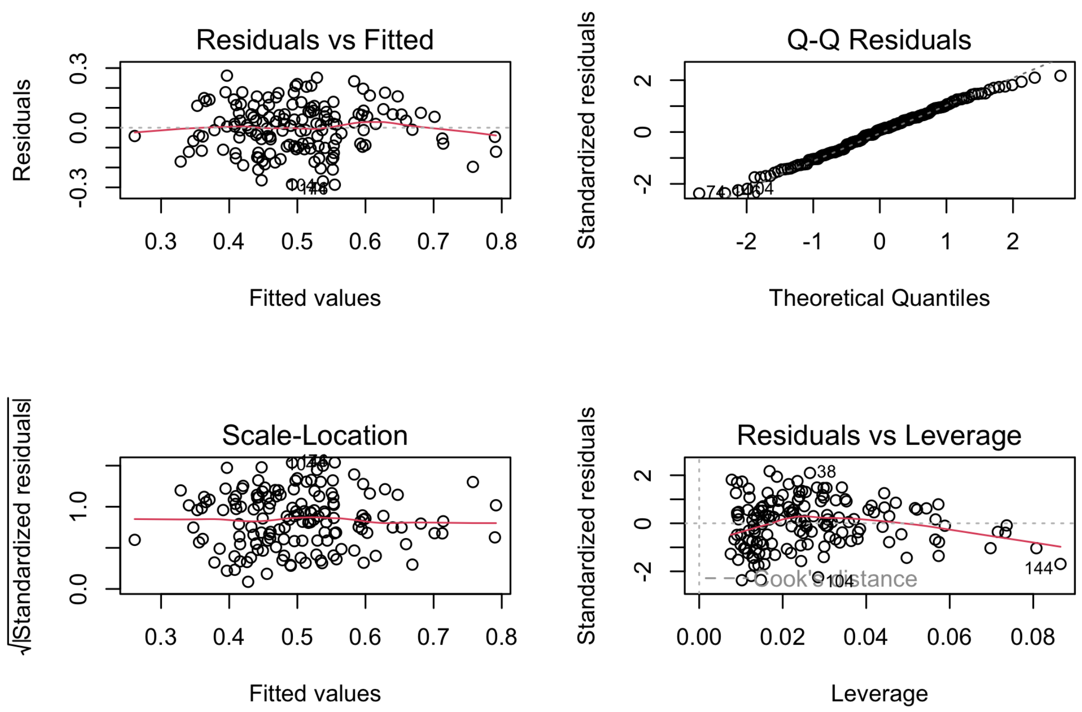

Figure 6: Comparison of Marginal Effects Between Periods.

5 Conclusion

References

Appendix