

Using Data Science Methods to Characterize Trends in the NBA

Anh-Minh Nguyen and Benjamin Stan
December 8, 2021

UNIVERSITY *of* WASHINGTON

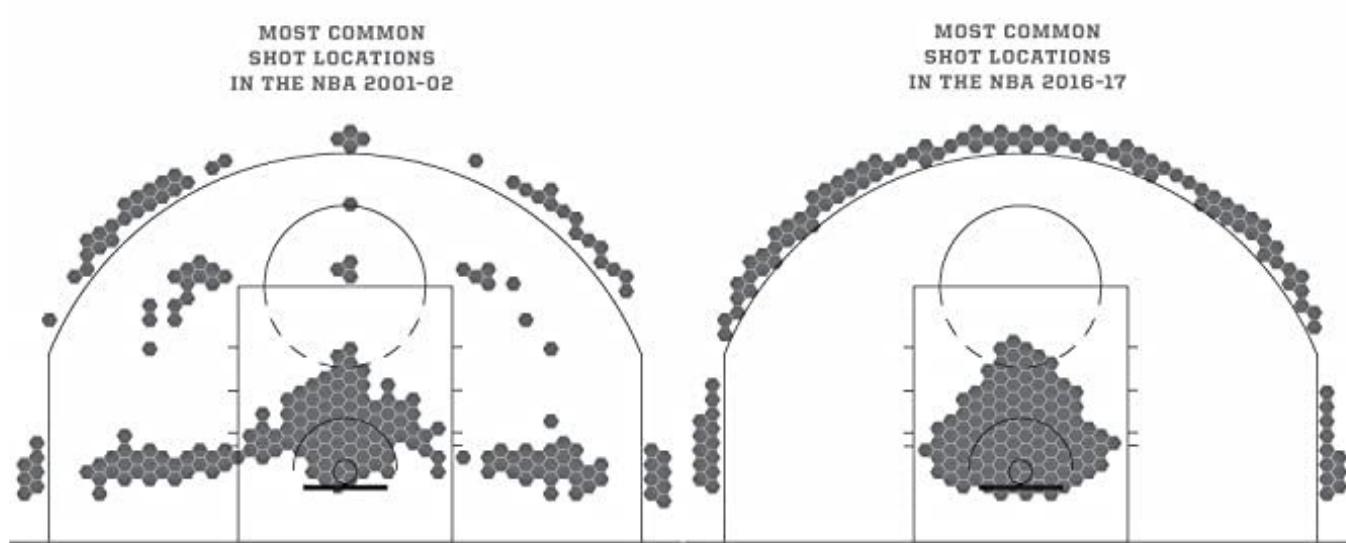


Background

UNIVERSITY *of* WASHINGTON



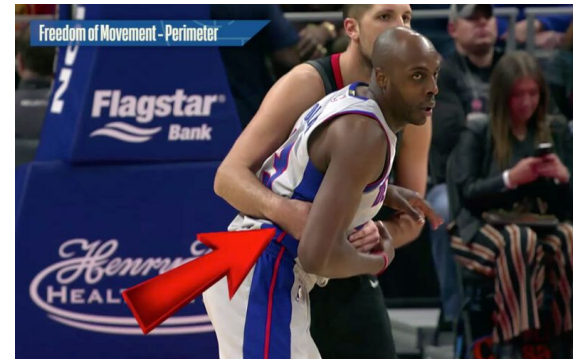
NBA gameplay is constantly evolving



UNIVERSITY *of* WASHINGTON

Effect of rule changes

- > Each year, NBA makes changes to rules or emphasizes/reinterprets existing rules
 - 2001: Implementation of zone defense and defensive three seconds violation
 - 2004: Removal of “hand-checking” reduces contact on the perimeter (top)
 - 2011: “Rip-through” moves no longer considered shooting fouls
 - 2018: “Freedom of movement” emphasis reduces contact by defenders (bottom)



Sources:

Johnson, D. (2018, November 21). NBA: How hand-check penalty changed basketball forever. Sportskeeda. Retrieved December 3, 2021, from <https://www.sportskeeda.com/basketball/how-hand-checking-foul-changed-the-nba-forever>
Calloway, L. (2018, October 12). New season brings new rules, a new official and new points of education. NBA.com. Retrieved December 3, 2021, from <https://www.nba.com/news/new-rules-official-points-education-2018-19>

NBA players dataset from Kaggle

- > 1996-2021
- > Includes season-long info
 - Basic stats (points, rebounds, assists) averaged per game
 - Player draft details (college, draft position)
- > True shooting percentage (TSP)
 - Defined using points (PTS) and true shot attempts (TSA).
 - > $TSA = (\text{field goal attempts}) + 0.44 * (\text{free throw attempts})$
 - > $TSP = PTS / (2 * TSA)$.

Aims and problem description

- > Does the rise of three-pointers lead to three point specialists and greater scoring depth for successful teams?
 - Compare the average PPG of non-leading scorers on championship teams against the average PPG of non-leading scorers on other teams
- > Are players becoming more effective at scoring?
 - Model the trend in true shooting percentage (TSP) from 1997 to 2021 to better understand the observed pattern
 - Interested in changepoints corresponding to rule changes: 2001, 2004, 2011, 2018

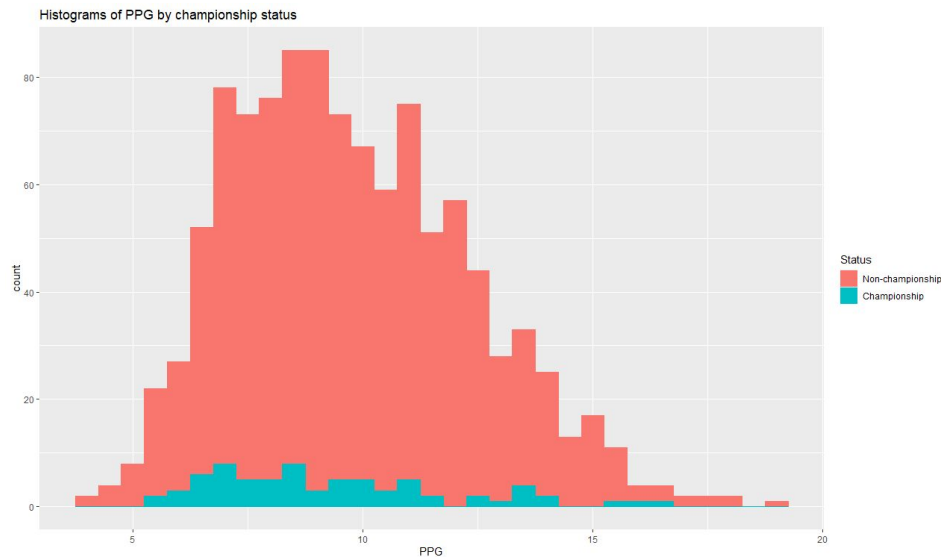
Evaluating Depth of Championship Teams

UNIVERSITY *of* WASHINGTON



Initiating the test

- > Identify the championship teams for each season (teams that played in the NBA finals for a specific season) [2015 - 2021]
- > Order the players by PPG on each team and consider only the 4th-9th ranked players
- > Calculate the mean PPG of players on championships teams and repeat for those on other teams
- > Take the difference between the two means

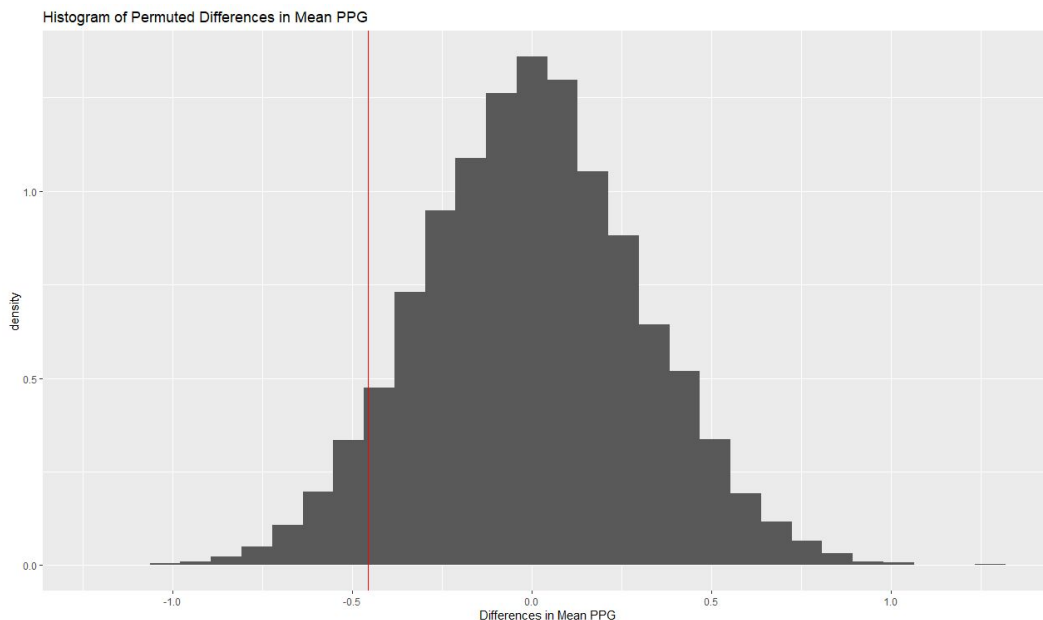


Original difference in means: -0.46

Developing the permutation algorithm

- > For first permutation, shuffle the label of whether or not a player was on a championship team, stratified by season
- > Re-calculate the mean PPG of those 4th-9th ranked players on championship teams and other teams
- > Take the difference between the two means and store it
- > Repeat previous steps 10,000 times in total
- > Compare original difference of means to array of permuted differences of means

Evaluating differences in mean PPG



Evaluate two-sided
p-value: 0.13

We fail to conclude
whether championship
teams have greater or
lesser depth in terms of
PPG than other teams.

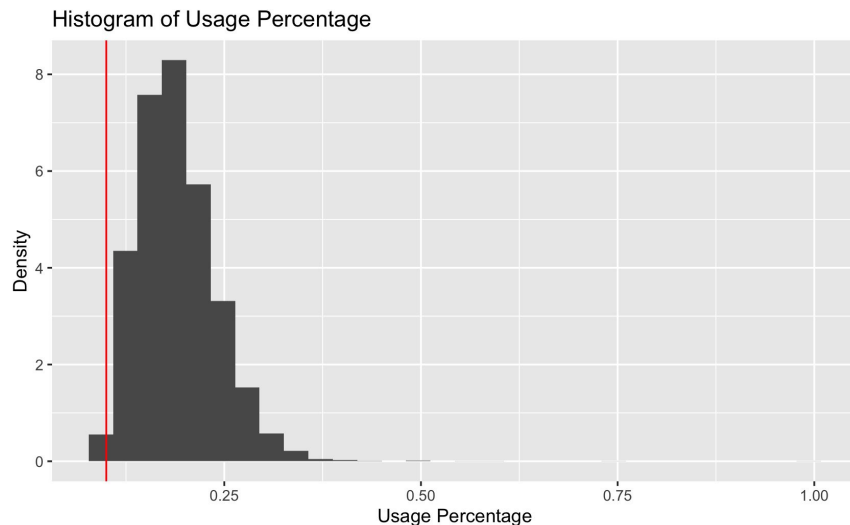
UNIVERSITY of WASHINGTON

Modeling True Shooting Percentage

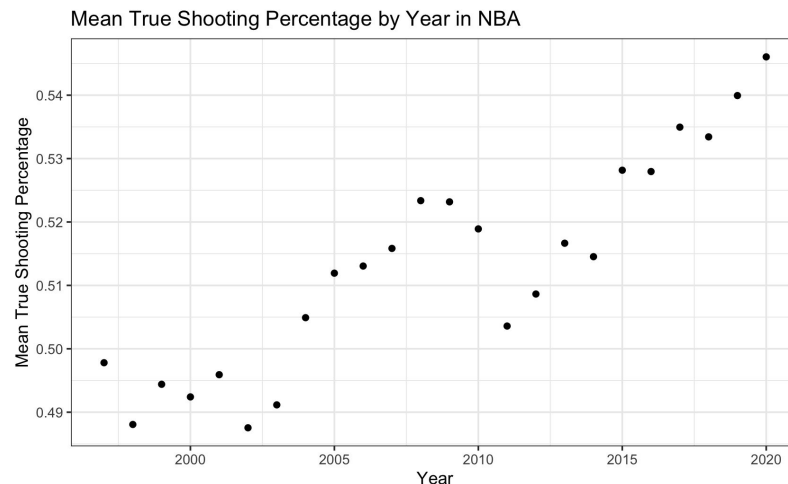
UNIVERSITY *of* WASHINGTON



Chose to model the mean TSP by year



Usage percentage: Estimate of the percentage of team plays used by a player while on the floor

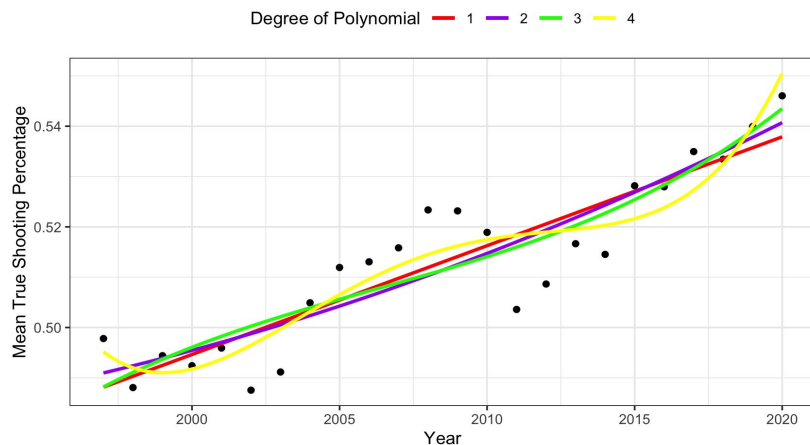


Considers players with usage percentage above 10%

UNIVERSITY of WASHINGTON

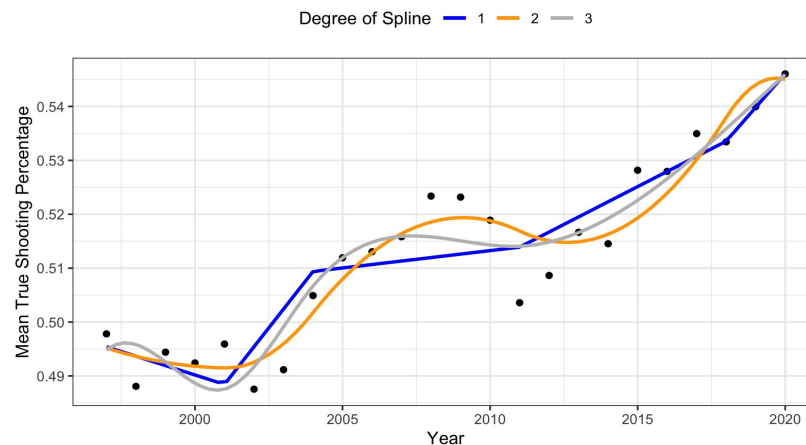
Visualization of model fits

Mean True Shooting Percentage by Year in NBA



Polynomials

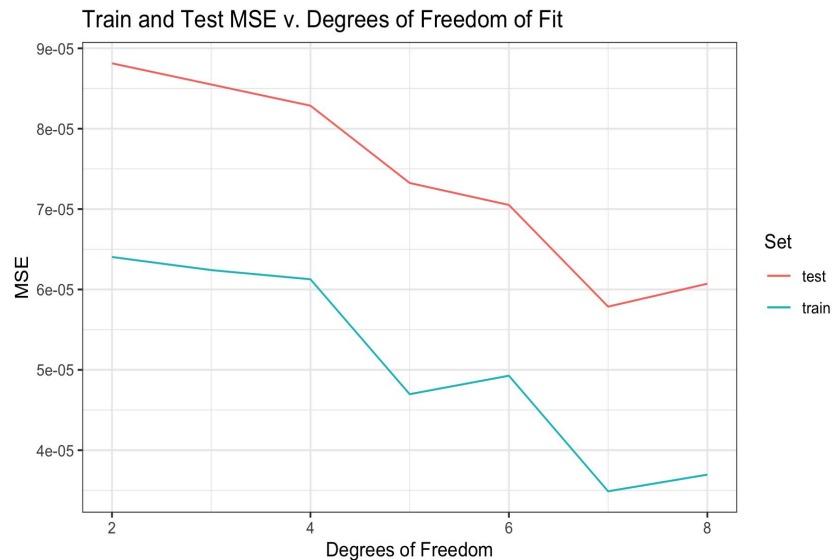
Mean True Shooting Percentage by Year in NBA



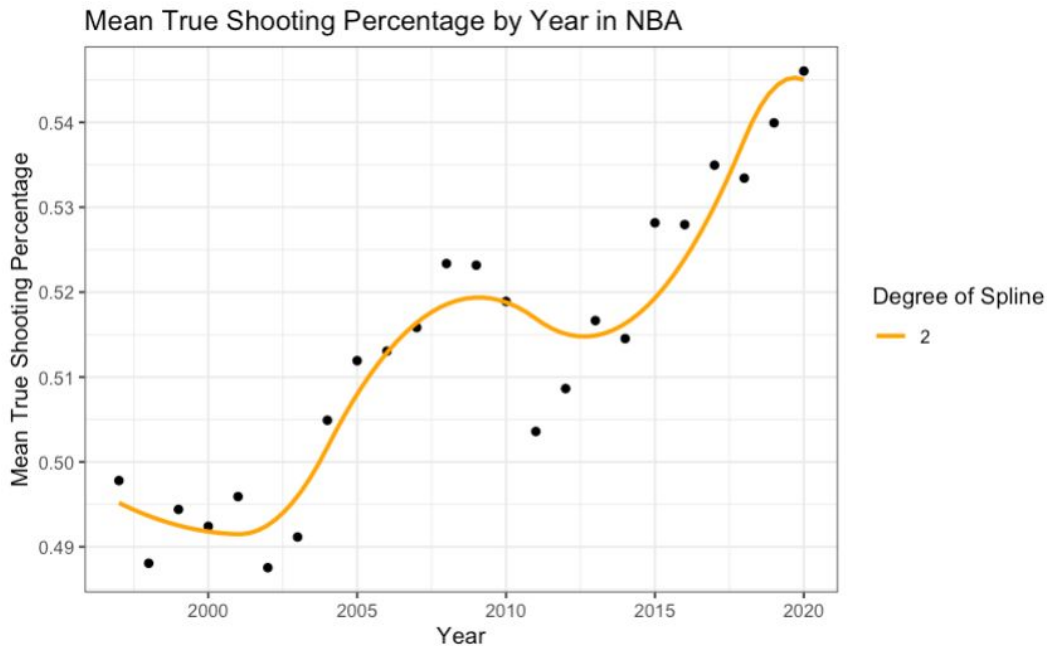
Splines

Determining the best model

- > Used 60/40 validation split, stratified by year
- > For polynomial (df 2-5)
 - $\text{df} = \text{order} + 1$
- > For splines (df 6-8)
 - $\text{df} = \text{order} + 5$
 - Four knots and intercept included
- > Slight increase in error from polynomial fit to spline (df 5 \rightarrow 6)
- > Quadratic spline (df 7) showed lowest test error

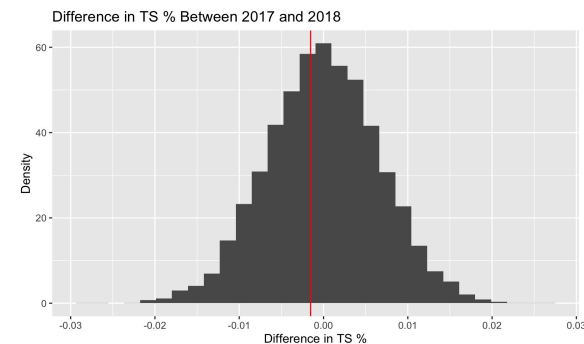
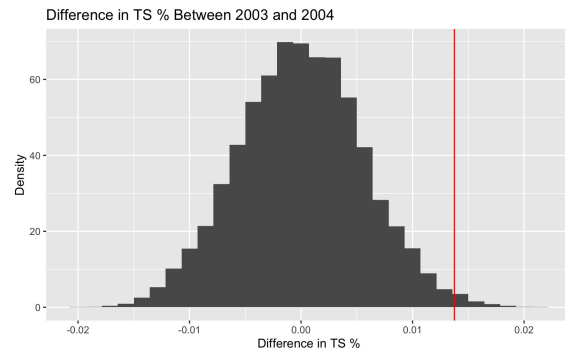


Visualization of model fit



Evaluating the effect of rule changes

Year	Rule Change	Point Est (after-before)	p-value
2001	Zone defense	0.35%↑	0.27
2004	Hand checking	1.38%↑	0.0085
2011	Rip-through move	1.53%↓	0.0011
2018	Freedom of movement	0.15%↓	0.41



Questions?



UNIVERSITY *of* WASHINGTON

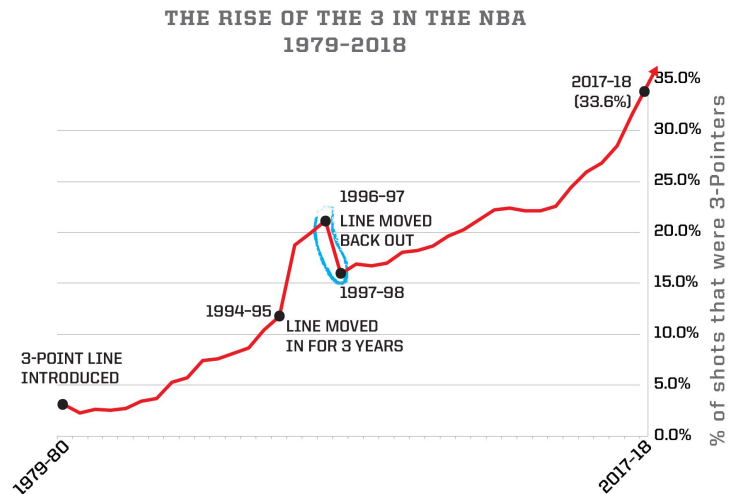
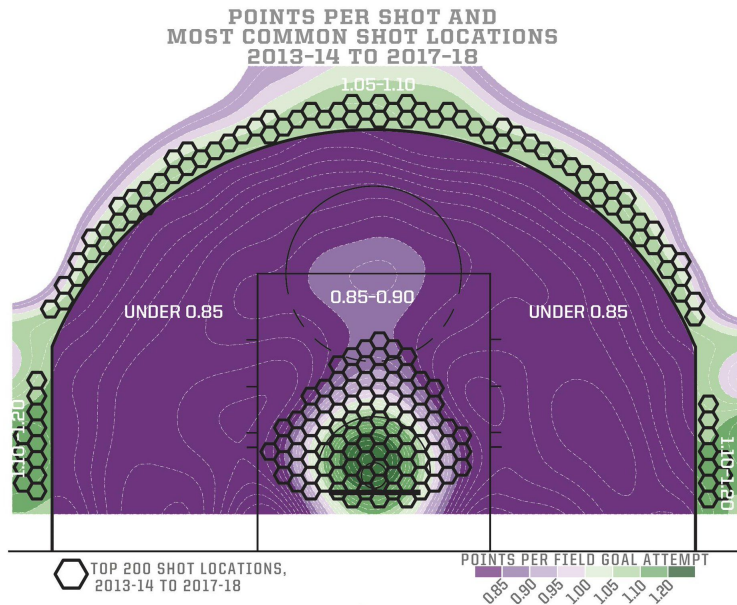


Appendix

UNIVERSITY *of* WASHINGTON



NBA gameplay is constantly evolving



UNIVERSITY of WASHINGTON