On Draft: The efficiency landscape in the NBA by draft round

BIOSTAT 653: *Final Project*

William Tackett, Jacob Vidergar

Dec 18, 2024

Background/Description

Over the past three or so decades, teams in the National Basketball Association (NBA) have evolved their game strategies and roster constructions for a few reasons. One is a focus on analytics, which has resulted in an emphasis on data driven strategies. For instance, the increased usage of the three point shot–long underlooked in game plans historically–indicates an emphasis on offensive efficiency. The surge in three point shots and overall offensive efficiency in the league is a well-characterized trend [5]. The effect of emphasizing the three point shot while simultaneously minimizing longer 2-point shots can be seen in a side-by-side comparison of shot charts from 2001-2002 and 2019-2020 (Figure 1). However, the effect of the NBA's uniquely dichotomous draft system, which focuses on two rounds, on this overall trend is less well known.

Figure 1: *Shot charts from the 2001-02 and 2019-20 season representing the top 200 most frequent shot locations. Taken directly from https://fadeawayworld.net*

Traditionally, players drafted in the first round have higher expectations and are more skilled and important to their teams than undrafted or second round players [4]. They end up playing more games, for example, compared to those drafted in Round 2 or who were undrafted [see Fig 2]. But how has the recent transformations in the NBA efficiency landscape affected this dynamic, if at all? There are other trends to consider, such as the increased presence of international talent in the league and a theoretical increase in talent domestically, both of which raise the floor of talent, potentially diminishing the talent/skill gap between first and second rounders.

As the NBA is an extremely well-documented organization both internally and externally, there is a lot of reliable and verified data available to investigate hypotheses. The dataset used in this study includes player performance metrics such as shooting efficiency (true shooting percentage, field goal percentage), usage percentage, points scored per game, alongside demographic as well as biometrics like weight and height from the 1996-97 season to the 2022-23 season, allowing us to conduct sub-analyses focusing on certain kinds of players like so-called big men (6’9” and above) who have changed in role over the course of the dataset [3].

The primary longitudinal outcome in this analysis will be true shooting percentage (TS%), a metric of shooting efficiency that accounts for two-point and three-point field goals as well as free throws [2]. By comparing the TS% trends for the two draft categories, we aim to identify patterns in shooting efficiency trends between the 1996-97 season and 2022-23 seasons.

In this analysis, we aim to investigate how shooting efficiency has changed over the last three decades and how this trend is affected by the factor of draft round. Specifically, does the historical discrepancy between first round and second round draft picks change over the time period of interest?

Methods

2.1 Inclusion Criteria

Because our dataset spans over 20 years of NBA players, it includes individuals who played in very few games. This resulted in these players having an inflated true shooting percentage because of their limited number of shot attempts. To address this, we restricted our analysis to players who participated in at least 50% of games each season.

2.2 Outcomes and Predictors

The main parameter of interest in our model is the true shooting percentage of each player. The true shooting percentage is a weighted measure of a player’s shooting efficiency that takes into account free throws, 2 point, and 3 point shots. Essentially, TS% takes half the points scored by a player and divides by the number of scoring changes. The 0.44 coefficient for FTA has been empirically determined and represents the percentage of free throws that are possession-ending (most free throws come in pairs so theoretically 50% would be possession-ending, but roughly 6% are not for various reasons–and 1s, offensive rebounds, etc).

Equation 1 (below) details the formula for the true shooting percentage of an individual player.

$$TS\% = \frac{PTS}{2(FGA + (0.44 \times FTA))}$$

Equation 1: True shooting percentage formula

Our outcome of interest was TS% as stated before. Other variables we considered were season, the principal time variable in this dataset, and draft round, our main regressor of interest, as well as height and weight. Age is of course an important factor in this analysis as players' ages are known to be related to their performance, potentially following a U-shaped curve [1]. However, age is highly related to season, of course, so the potential for collinearity issues arises. To balance these two considerations, we created a variable called career stage, a three level factor dividing players into rookie, mid-career, and veteran status. As described in the model diagnostics section, this new variable does not suffer from collinearity issues.

2.3 Models

Because our dataset contained repeated observations of the same players over the course of their careers, we elected for a mixed effect model to model our data. This type of model accounts for the fact that observations of the same player from different seasons are not independent of each other, as well as the individual differences that affect their performance which are not captured by the fixed effects. Including player as random effect also enables us to to estimate how much of the total variance is due to between-player differences versus within-player differences over time.

For the fixed effects in our model, we included:

- **Season • Draft Round:** A term representing the interaction between the current league year (1996-2023) and the round each player was taken in the draft (categorized into round 1 or round 2/undrafted)
- **Height:** The height of each player (cm)
- **Weight:** The weight of each player (kg)
- **Career Stage:** Three categories based on player age
    - Rookie (age < 25)
    - Mid-Career (25 ≤ age ≤ 30)
    - Veteran (age > 30)

Random effects:

- **Player Name:** The name of each player was chosen as a random effect to account for the individual differences between each player.

Model 1: TS% ~ $\beta_0 + \beta_1(\text{Season*Draft Round}) + \beta_2\text{Height} + \beta_3\text{Weight} + \beta_4\text{Career\_Stage} + u_{\text{Name}} + \varepsilon$

Multiple other models were implemented to investigate potential nonlinearity effects of height and weight, namely a model (Model 2) with height and weight as squared terms and a spline-based model (Model 3).

Model 2: TS% ~ $\beta_0 + \beta_1$(Season*Draft Round) + β_2Height^2 + β_3Weight^2 + β_4Career_Stage + u_{Name} + ε

Model 3: TS% ~ $\beta_0 + \beta_1$(Season*Draft Round) + splines(Height) + splines(Weight) + β_4Career_Stage + u_{Name} + ε

Results

Preliminary plots of the data showed evidence of a general increase in TS% overall as the seasons progressed from 199-97 to 2022-23 (Figure 2). Additionally, breaking this trend down by draft round indicates a potential shift in the gap between first and second rounders, at least in terms of skill set (Figure 2). Plotting height and weight vs TS% also offered early insight into that relationship, suggesting as players get heavier and taller, their TS% potentially gets higher.

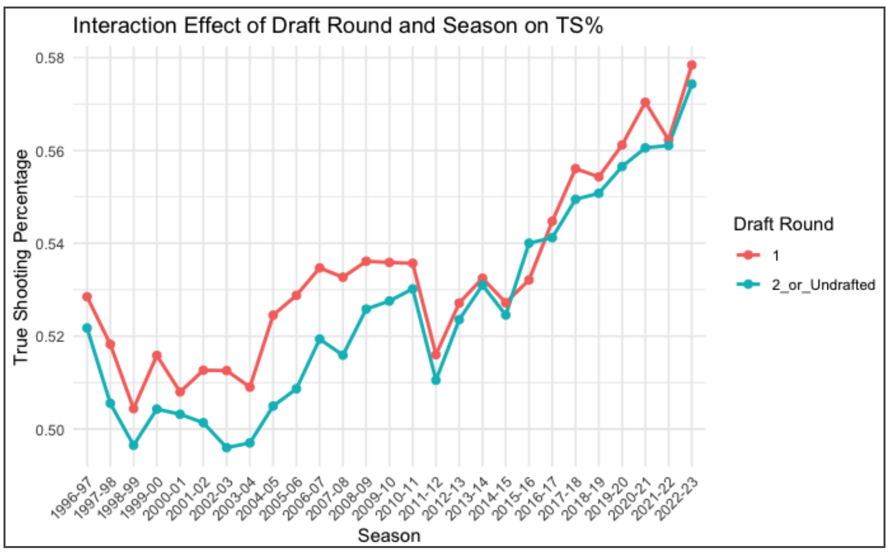

Figure 2: Spaghetti plot of TS% by draft round

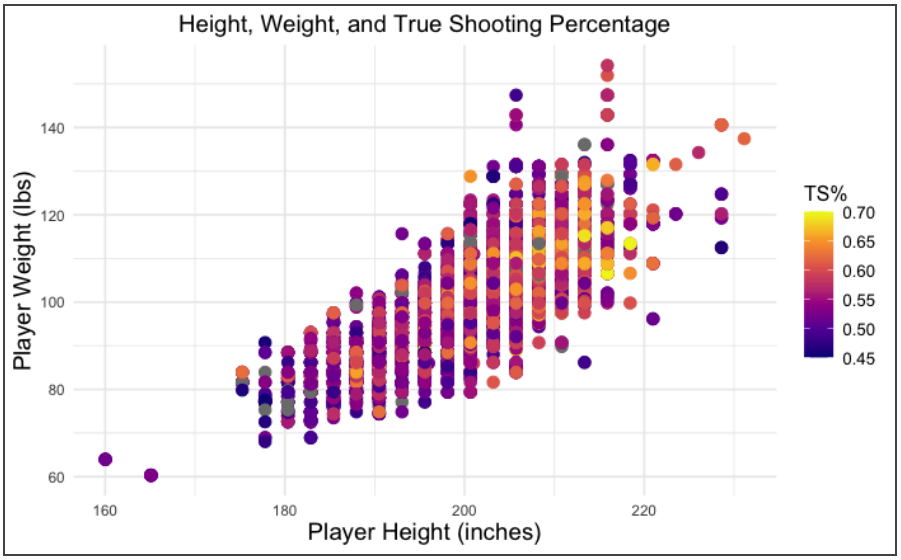

Figure 3: Heatmap of player height and weight vs TS%

Results from Model 1 are summarized in Table 1.

Predictors	*Estimates*	*std. Error*	*CI*	*p*
(Intercept)	-3.65	0.24	-4.12 – -3.18	**<0.001**
season continuous	0.00	0.00	0.00 – 0.00	**<0.001**
draft round combined new [2_or_Undrafted]	-0.40	0.38	-1.14 – 0.35	0.298
player height	0.00	0.00	0.00 – 0.00	**<0.001**
player weight	0.00	0.00	0.00 – 0.00	**0.009**
career stage [Mid-Career]	0.01	0.00	0.01 – 0.01	**<0.001**
career stage [Veteran]	-0.00	0.00	-0.01 – -0.00	**0.001**
season continuous × draft round combined new [2_or_Undrafted]	0.00	0.00	-0.00 – 0.00	0.303

Table 1: Regression table for Model 1

For the investigation of nonlinear effects of height and weight using Model 2, plots of predicted values for height and weight based on the model were generated:

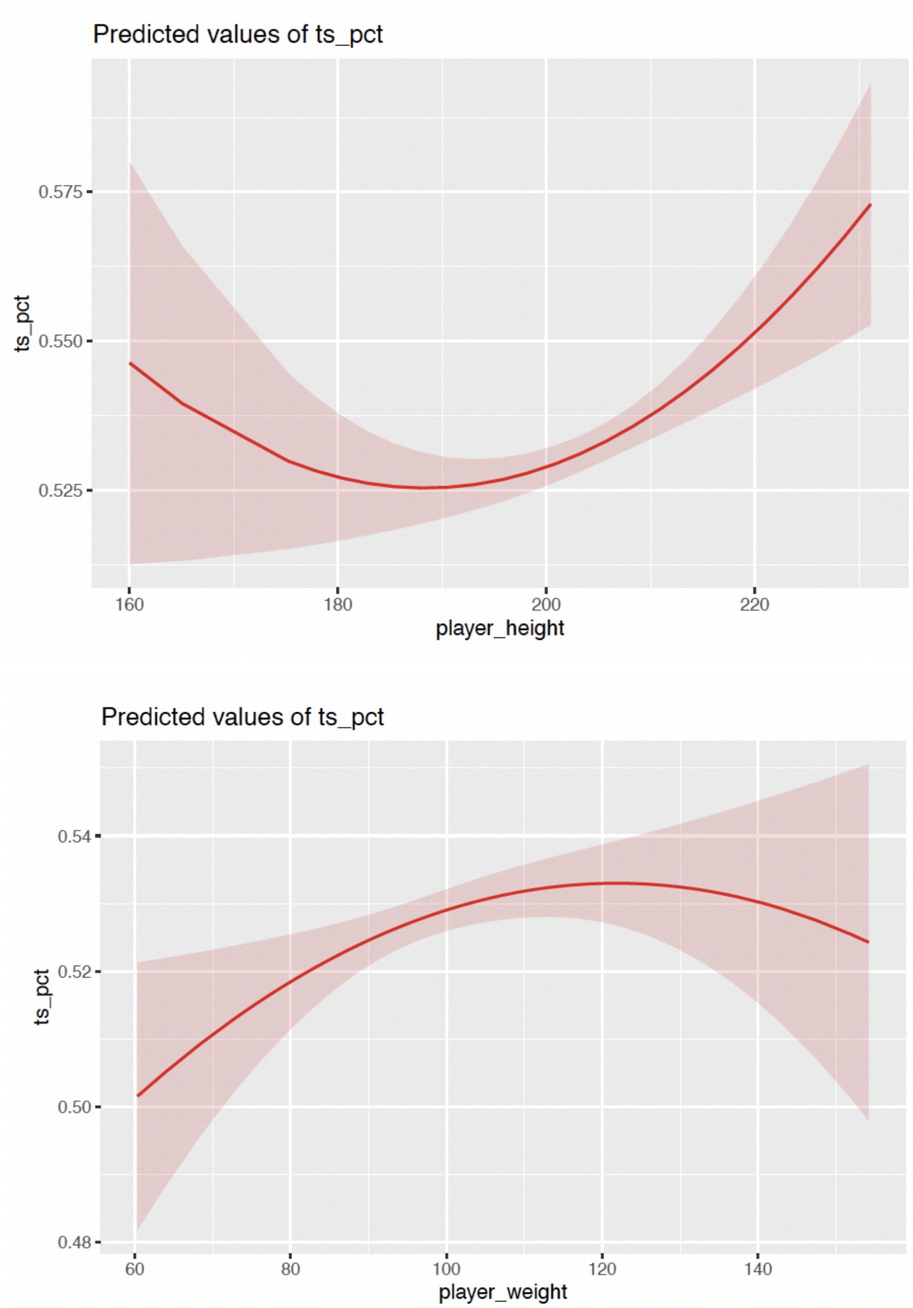

Figure 4: Predicted values for height (top) and weight (bottom) from Model 2

To assess the different model approaches, a likelihood ratio test was performed for Model 1, 2, 3. The results are summarized in Table 2.

	df	AIC
Model 1	10	-30703.06
Model 2 (quadratic)	12	-30727.90
Model 3 (spline)	14	-30693.14

Table 2: Results from likelihood ratio test of considered models

Diagnostics

All diagnostic plots from model 1.

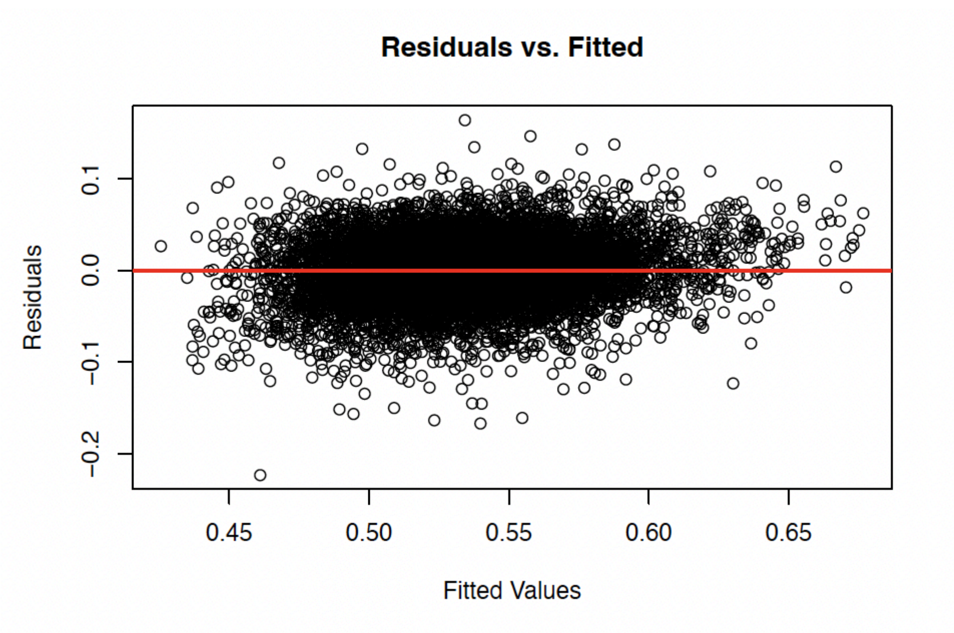

Figure 5: Residuals vs fitted values for Model 1

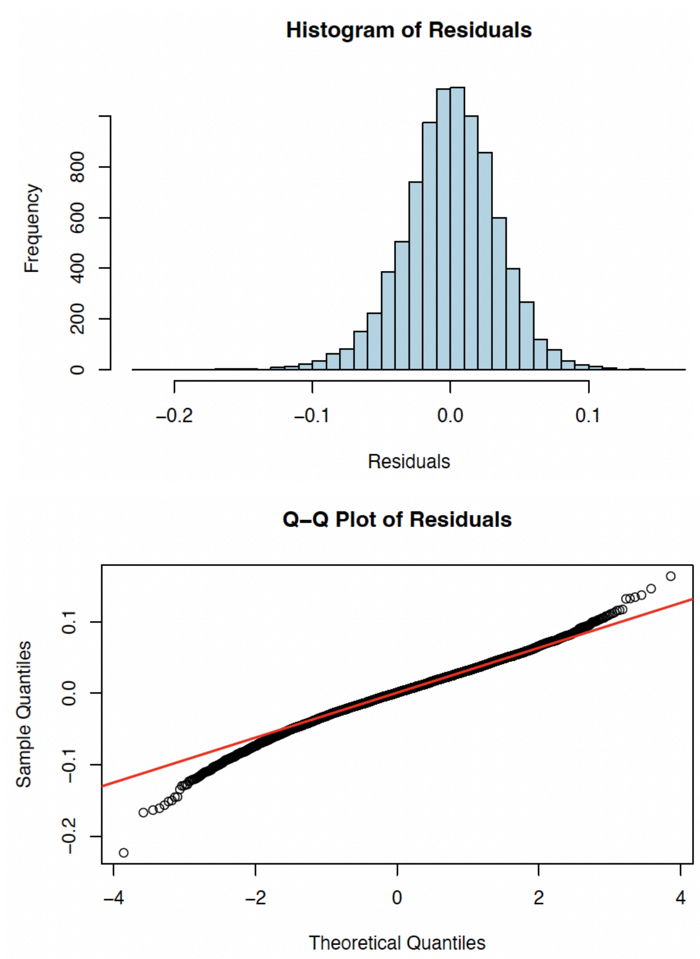

Figure 6: Histogram (top) and QQ-plot (bottom) for Model 1.

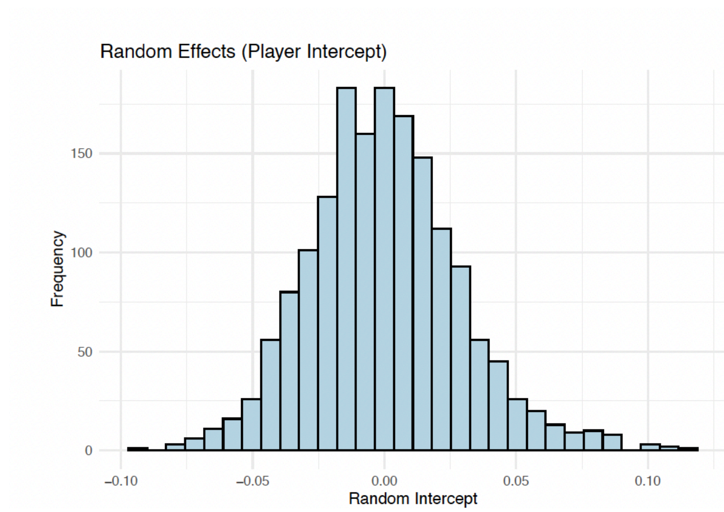

Figure 7: Histogram of random effects (player intercept) for Model 1

Discussion

This study investigated trends in true shooting percentage (TS%) among NBA players over the past three decades, with a focus on how draft round status (first round versus second round/undrafted) interacts with these trends. Our findings provide insight into the evolving efficiency landscape in the NBA and highlight the importance of considering player-specific and contextual factors when evaluating shooting efficiency.

As expected, our results show that TS% has increased steadily over the past three decades, consistent with the broader league trend of prioritizing offensive efficiency and three-point shooting. This increase is likely influenced by shifts in game strategy, training methodologies, and the overall rise in talent levels. The interaction term between season and draft round status did not achieve statistical significance in our primary model. Thus, our hypothesis that first and second rounders have converged in efficiency since 1996 was not substantiated. However, this is just one metric, and further studies should look at other measures of efficiency and player effectiveness.

Quadratic and natural spline models revealed nonlinear relationships between height/weight and TS%, indicating that taller and heavier players tend to have higher shooting efficiency, but with diminishing returns at extreme values. These results align with the evolving roles of taller players, who have increasingly developed perimeter shooting skills to complement traditional post-play dominance.

We also found that the TS% trends differed across career stages, with mid-career players achieving the highest efficiency, followed by rookies and then veterans. This U-shaped relationship aligns with previous findings suggesting that physical decline and role changes (e.g., fewer shot attempts) may reduce efficiency in the later stages of a player's career.

Because our outcome variable varies from 0 to 1, we considered using a transformation. We tried a logit transformation, and found that the original model was a significantly better fit based on AIC.

One limitation of our dataset was the small number of variables to work with. Factors such as injury, coaching, individual team strategy all play a role in our outcome variable but were not controlled for in this model. The flip side is that this dataset contains all relevant data for the population of interest and is thus not technically a random sample. Additionally, because the NBA is so well documented, there was no missing data in the dataset whatsoever.

This study provided a rigorous analysis of longitudinal trends in the NBA and was able to support beliefs about draft round, career stage, and size on the performance of a player, as well as efficiency trends happening in the league as a whole. One interesting potential finding was the apparent convergence in efficiency of first and second rounders right around the time of the lock-out season (Figure 2), an important season in the league historically as it was only partially completed due to players striking and ongoing CBA talks. Perhaps rule changes or something else caused this effect or perhaps it is an artifact, but it warrants future research.

References

1. Brady, B. (2017, November 17). What Happens to NBA Players When They Age? [Review of What Happens to NBA Players When They Age?]. *Harvard Sports Analysis Collective*. https://harvardsportsanalysis.org/2017/11/what-happens-to-nba-players-when-they-age/
2. Jacobs, J. (2019, January 9). True Shooting Percentage Part I: Introduction and Framework for Advancement [Review of *True Shooting Percentage Part I: Introduction and Framework for Advancement*]. *Squared Statistics: Understanding Basketball Analytics*. https://squared2020.com/2019/01/09/true-shooting-percentage-part-i-introduction-and-framework-for-advancement/
3. *NBA Players*. (n.d.). Www.kaggle.com. https://www.kaggle.com/datasets/justinas/nba-players-data/data
4. *Statistical Analysis Primer*. (n.d.). Www.nba.com. https://www.nba.com/thunder/news/stats101.html
5. Zając, T., Mikołajec, K., Chmura, P., Konefał, M., Krzysztofik, M., & Makar, P. (2023). Long-Term Trends in Shooting Performance in the NBA: An Analysis of Two- and Three-Point Shooting across 40 Consecutive Seasons. *International journal of environmental research and public health*, *20*(3), 1924. https://doi.org/10.3390/ijerph20031924