Exploring various NBA draft value curves

Joseph Bewtra, Ryan S. Brill

November 25, 2024

Abstract

NBA teams often trade draft picks. We are interested in the relative value of draft positions, which determines whether a team should accept or reject a trade, or which draft picks it should offer in a trade to make it fair. In this work, we explore various NBA draft value curves. We introduce a novel method, mapping player performance measures (e.g., WAR, RAPTOR, BPM) to salary using Gamma regression in order to constrain draft value curves to be positive. We find that, depending on the measure of performance value and the method of aggregation (e.g., mean or median), draft value curves are wildly different.

1 Introduction

NBA teams often find themselves wanting to trade draft picks. A general manager may want to trade up to draft a particular player before he is taken by another team. He may also want to trade up to be able to draft a player of a certain caliber at a particular position. In exchange for trading up, a general manager often offers a trade consisting of current and/or future draft picks. This naturally leads to the following questions. If a team wants to trade for a particular draft pick, which picks should it offer in exchange to ensure they make a good trade? And if a team is offered a bundle of draft picks in exchange for another bundle of draft picks, should it accept or reject the trade? We are interested in the relative value of draft picks.

Suppose you are an NBA general manager who is offered a bundle of draft picks in exchange for a bundle of your draft picks. Suppose prior to the draft an omniscient oracle told you who will get drafted with each pick and how "good" they will be, or how much "value" their performance will provide to your team. Then you should accept the trade if the "value" of the picks offered to you exceeds that of your original picks. This toy scenario suggests the following questions. First, how should we quantify the value of a player's performance? Second, how can we account for the fact that we do not have access to a drafted player's future performance?

1.1 Measuring player performance

In basketball analytics, there are many "all-in-one" measures of player impact that capture player performance in a way that measures all players on the same scale.[1,2] In this work, we consider three such measures: BPM, RAPTOR, and WAR. Box Plus Minus (BPM) uses a player's box score (game summary statistics including points, rebounds, assists, etc.) to estimate his plus minus (the difference in his team's point differential when he is on the court versus off the court).[3] RAPTOR, which stands for Robust Algorithm (using) Player Tracking (and) On/Off Ratings, is a plus-minus metric that incorporates the box score and other player summary statistics derived from player tracking and play-by-play data.[4] Finally, WAR, which stands for Wins Above Replacement, estimates how many fewer wins a player's team would have had if a replacement-level player had played instead of him.[5]

We begin with each of these three measures, BPM, RAPTOR, and WAR, on the player-season level. We then compute six measures that encapsulate the career performance of each player – BPM, RAPTOR, and WAR per season and total career BPM, RAPTOR, and WAR.

1.2 Traditional NFL draft trade value charts

Draft trade value charts were originally created for the NFL draft. At the request of Dallas Cowboys coach Jimmy Johnson, vice president Mike McCoy created the first draft value chart based on gut instinct and past trades, known today as the Jimmy Johnson chart.

[1] https://hoopshype.com/lists/advanced-stats-nba-real-plus-minus-rapm-win-shares-analytics/
[2] https://www.cryptbeam.com/2021/05/21/the-10-best-nba-impact-metrics/
[3] https://www.basketball-reference.com/about/bpm2.html/
[4] https://fivethirtyeight.com/features/introducing-raptor-our-new-metric-for-the-modern-nba/
[5] https://www.nba.com/stats/players/advanced-leaders

Massey and Thaler (2013) later created an NFL draft trade value curve using statistical methods. Their thinking on a high level proceeds as follows. Prior to the draft, we do not know how well a player drafted at position x will perform in the NFL. Thus, we think of the performance outcome Y associated with pick x as a random variable, denoted by a capital letter. From the dataset $\{(x_i, Y_i)\}$ of all recent draft picks in NFL history, we want to learn the de-noised relationship between pick number x and performance outcome Y. Traditional analyses use a standard approach but with differing measures of first contract performance (Massey and Thaler, 2013; Stuart, 2012; Fitzgerald and Spielberger, 2024; Pro Football Focus, 2024; Baldwin, 2024). The de-noised relationship they estimate from data is $x \mapsto \mathbb{E}[Y|x]$, the expected value of Y given x. These are known as performance value curves.

In this study, we adapt this approach to the NBA draft. In particular, thinking of an NBA player's performance outcome Y as a random variable and letting the realization of Y be his total career BPM, RAPTOR, or WAR or his BPM, RAPTOR, or WAR per season, we consider draft curves of the form $x \mapsto \mathbb{E}[Y|x]$ and $x \mapsto \text{Median}(Y|x)$. We find that, depending on the measure of performance value and the method of aggregation (e.g., mean or median), draft value curves are wildly different.

Our work differs from previous research in a few ways. We explore additional outcome variables, we explore both mean and median curves (traditional curves use just the mean or expected value), and we plot multiple curves on top of each other to compare and contrast them (previous work visualize draft value curves separately from each other, masking how different they can be depending on the methodology used). Tony ElHabr[6] wrote a good overview of existing NBA draft value curves[7] on his blog. He discusses six approaches analysts have used to estimate the relationship between draft position and basketball production, from Justin Kubatko, Aaron Barzilai, Arturo Galletti, Nate Silver, Saurabh Rane, and Michael Lopez. These methodologies differ in their choices of player performance outcome Y (e.g. WS, VORP, etc.), sample size (i.e range of draft years), the time span for evaluating production (e.g. four years, career, etc.), whether to account for the cost of

[6] `https://tonyelhabr.rbind.io`

[7] `https://tonyelhabr.github.io/nba-decision_analysis/what-research-says-about-nba-draft-pick-value.html`

a draft pick (i.e., to examine surplus value or to not), and their regression methods (e.g. linear-log, LOESS, etc.). These analysts arrived at a similar qualitative conclusion: earlier draft picks are worth more than later draft picks. But it is not clear the extent to which these draft curves differ and we haven't seen a thorough comparison of each curve's outputs or recommendations.

2 Methods

In this section, we compute various draft value curves.

2.1 Data and code

To construct an NBA draft value curve, we create a dataset in which each row is a drafted NBA player. For each drafted player from 2010 to 2021, we record several variables. We record his draft position in $\{1, ..., 60\}$, his total career BPM, RAPTOR, and WAR (referred to henceforth by Total BPM, RAPTOR, or WAR), and his career BPM, RAPTOR, and WAR divided by the number of years he played (referred to henceforth by Average BPM, RAPTOR, or WAR). We also record his total career salary and his career salary divided by the number of years he played. We record salary in each year not in terms of dollars, but as a proportion of the highest super-max contract a player could receive that year, to adjust for inflation. Our dataset, and the code for this study, is publicly available at `https://github.com/snoopryan123/NBA_draft_curves_Joey`.

2.2 Initial draft value curves

For concreteness, let's construct an expected performance value curve that values a player's performance using his career Total BPM. First, for each draft position $x \in \{1, ..., 60\}$, we compute the Total BPM averaged across all players drafted at pick x in our dataset. These empirical averages are the black dots in the "mean Total BPM" panel of Figure 1 (the graph in the upper right corner). Then, we smooth these empirical values by fitting a spline. This produces the smooth curve in that panel (the teal curve). Constructing curves with other performance measures (see Section 1.1) and other aggregators (e.g., median rather than expected value) is a similar process. We visualize these each of these curves in Figure 1.

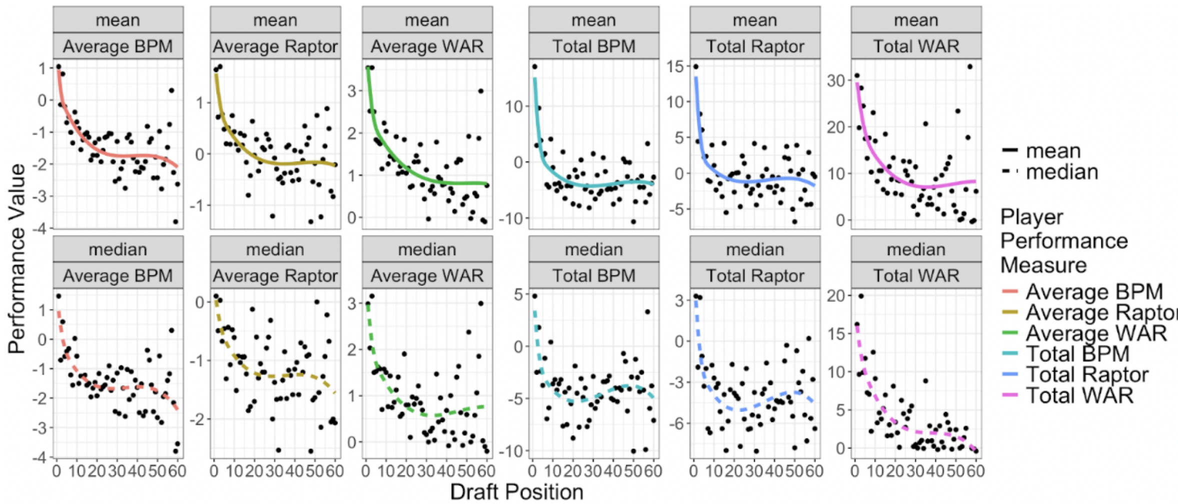

Figure 1: Mean or median (linetype) player performance value (y axis) as a function of draft position (x axis) for each player performance measure (color). We display twelve draft value curves, one for each combination of the six performance measures (Total BPM, RAPTOR, and WAR and Average BPM, RAPTOR, and WAR) and two de-noisers (mean and median).

Each of the draft value curves in Figure 1 are on a different scale. To compare them and to use them to make NBA draft trades, we normalize them to be on the scale. As in Massey and Thaler (2013), we normalize by dividing by the value of the first pick. The value of the first pick becomes one and the value of pick x becomes its value relative to the first pick. We visualize the normalized draft value curves in Figure 2.

These curves are now on the same scale, but they have some problems. First, four of the curves have negative values, which doesn't make sense. We want the value of pick x relative to the first pick to be some fraction $v(x) \in [0, 1]$, suggesting that $1/v(x)$ of that pick has the same value as the first pick. For example, the value of the twentieth pick relative to the first pick according to the Mean Average BPM curve is 1/2, suggesting that two twentieth picks are worth one first pick. This logic breaks down when pick x has a negative value. For example, the value of the seventh pick relative to the first pick according to the

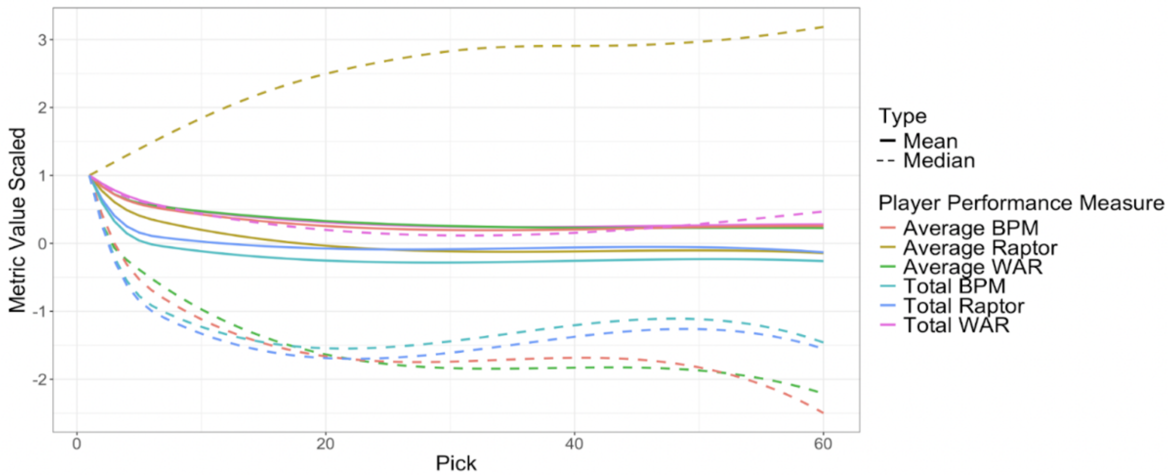

Figure 2: Mean or median (linetype) value relative to the first pick (y axis) as a function of draft position (x axis) for each player performance measure (color). We display twelve normalized draft value curves, each relative to the first pick, one for each combination of the six performance measures (Total BPM, RAPTOR, and WAR and Average BPM, RAPTOR, and WAR) and two de-noisers (mean and median).

Median Total BPM curve is -1, so it is not clear how many seventh picks have the same value as the first pick.

Second, the Median Average RAPTOR curve is monotonically increasing, which doesn't make sense. This happens because the median average RAPTOR for the first pick, and for most draft positions, is negative. We want each draft value curve to be monotonically decreasing, as we expect earlier picks to be more valuable than later picks.

2.3 Mapping player performance to salary

To fix the issues with the NBA draft value curves we constructed in the previous section, we map player performance to salary (adjusted for inflation). Salary is a non-negative measure of player value, eliminating the problems arising from negative values.

Crucially, we do not measure player performance using his actual average or total career salary because the NBA imposes a maximum salary for any individual contract, which conceals the true performance value of players. For example, in the 2023-24 season both

Nikola Jokić and Bradley Beal had max contracts,[8] but the former contributed much more to winning basketball games than the latter (Jokić's BPM was 13.2 and Beal's was 0.4).[9]

Instead, we map each player performance measure (Average BPM, Average RAPTOR, Average WAR, Total BPM, Total RAPTOR, and Total WAR) to salary. We construct a monotonic mapping so that an increase in performance value is associated with an increase in predicted salary. This prevents the truncation in value caused by the maximum salary. We map average annual performance (e.g., Average BPM, RAPTOR, or WAR) to average annual salary and total career performance (e.g., Total BPM, RAPTOR, or WAR) to total career salary. To adjust for inflation, we measure salary as a proportion of the maximum contract allowed for a single player that year.

For concreteness, let's consider mapping total career BPM, denoted y, to total career salary (as a proportion of the max contract), denoted S. To do so, we use Gamma regression (Bossio and Cuervo, 2015). Thinking of the salary S associated with a player having performance value y as a random variable – players with the same total BPM could end up having differing career salaries – we want to predict salary from performance (i.e., estimate $\mathbb{E}[S|y]$) in a way that aligns with this distribution. Gamma regression is a good choice to estimate $\mathbb{E}[S|y]$ because the distribution of S is non-negative and has a right tail, which changes as performance y changes (in particular, average salary $\mathbb{E}[S|y]$ increases as performance y increases).

The Gamma regression modeling process proceeds as follows. For each player i, we model $S_i|y_i \sim \text{Gamma}(k(y_i), \theta(y_i))$. The shape parameter $k = k(y_i)$ and the scale parameter $\theta = \theta(y_i)$ vary with y_i. The conditional mean is $\mu(y_i) := \mathbb{E}[S_i|y_i] = k(y_i) \cdot \theta(y_i)$ and the conditional variance is $\sigma^2(y_i) := \text{Var}(S_i|y_i) = k(y_i) \cdot \theta^2(y_i)$. To model how the distribution of S changes as y changes, we impose a linear model, $\mu(y_i) = e^{\beta_0 + \beta_1 \cdot y_i}$ and $\sigma^2(y_i) = \beta_2$. The $\exp(\cdot)$ link function constrains predicted salary to be positive, which is crucial to ameliorate the problems from Section 2.2. From our dataset, we estimate the parameters β_0, β_1, and β_2 using maximum likelihood.[10] Predicted salary given performance y_i is then $\hat{S}_i = e^{\hat{\beta}_0 + \hat{\beta}_1 \cdot y_i}$.

[8] https://www.espn.com/nba/salaries/_/year/2024/seasontype/4

[9] https://www.basketball-reference.com/leagues/NBA_2024_advanced.html

[10] https://cran.r-project.org/web/packages/Gammareg/index.html

2.4 Improved draft value curves

In the previous section we mapped each player performance measure to salary. Now, we construct draft value curves using predicted salary as a measure player performance. In Figure 3 we visualize pre-normalized draft value curves using our twelve measures of player performance mapped to salary. As desired, each of the curves are non-negative, and the curves are nearly entirely monotonically decreasing.

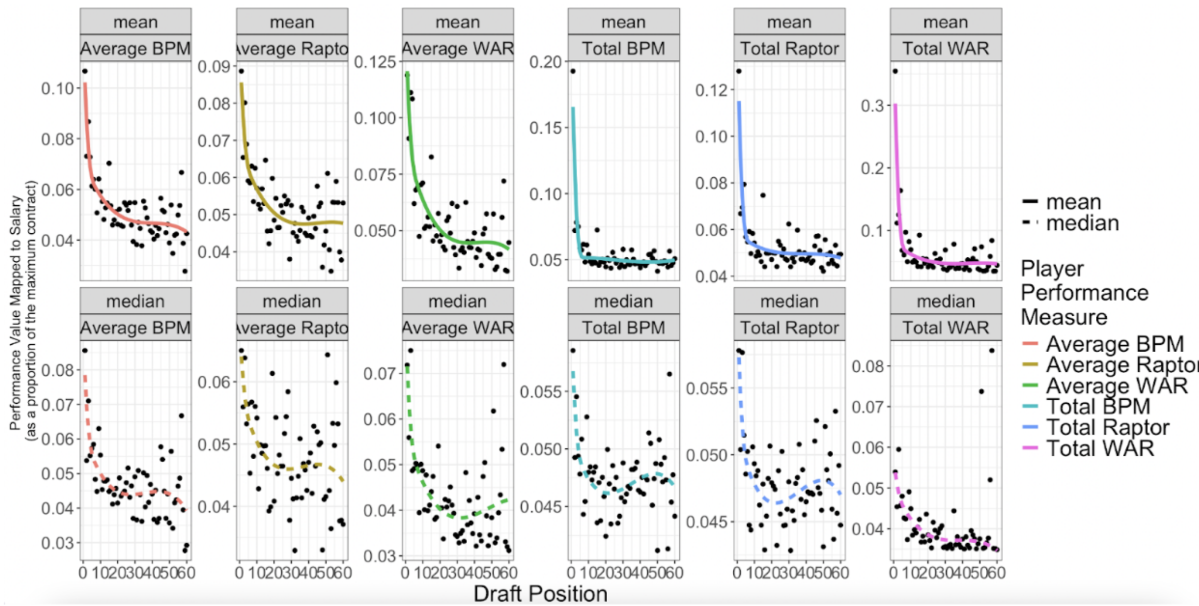

Figure 3: Mean or median (linetype) player performance value mapped to salary as a proportion of the max contract (y axis) as a function of draft position (x axis) for each player performance measure (color). We display twelve draft value curves, one for each combination of the six performance measures mapped to salary (Total BPM, RAPTOR, and WAR and Average BPM, RAPTOR, and WAR mapped to salary) and two de-noisers (mean and median).

Then, we normalize each curve so that the value associated with pick x is its value relative to the first pick. We visualize the normalized draft value curves in Figure 4. This plot is fascinating for several reasons. First, each median curve lies far below its mean curve. We suspect this is because lottery picks have a higher concentration of extremely good outlier players whose value doesn't affect the median as much as the mean. Second, the twelve draft value curves are wildly different. For example, according to median Total BPM curve, the 30^{th} pick is worth 0.82 of the first pick, but according to the mean Total WAR curve it is worth just 0.16 of the first pick.

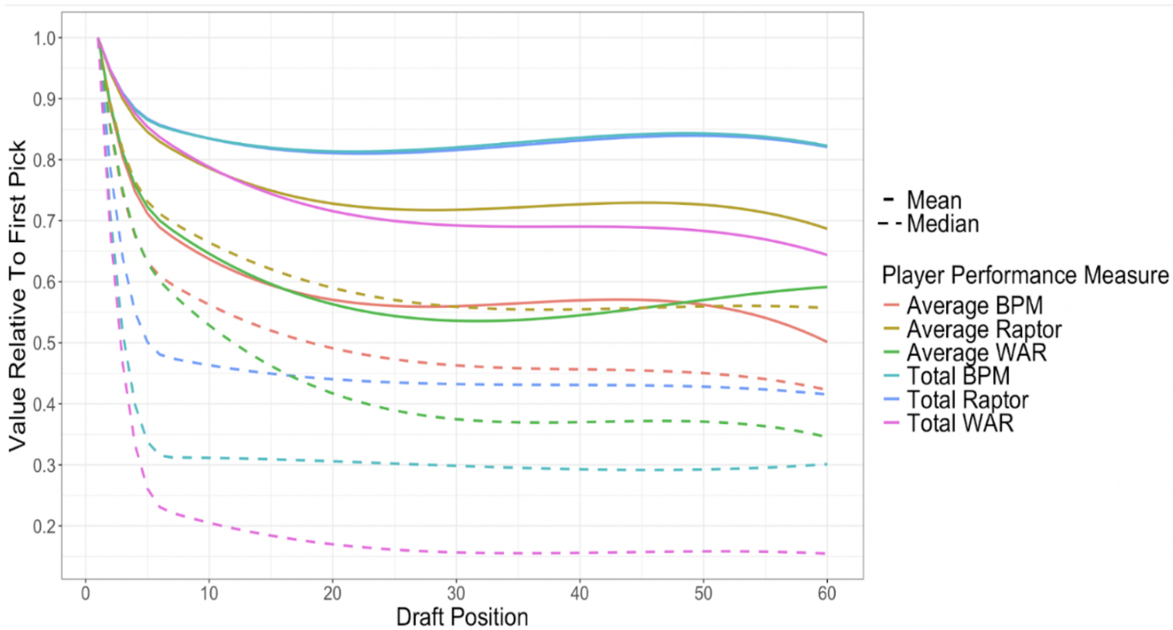

Figure 4: Mean or median (linetype) value relative to the first pick (y axis) as a function of draft position (x axis) for each player performance measure mapped to salary (color). We display twelve normalized draft value curves, each relative to the first pick, one for each combination of the six performance measures mapped to salary (Total BPM, RAPTOR, and WAR and Average BPM, RAPTOR, and WAR mapped to salary) and two de-noisers (mean and median).

It is not immediately clear which draft value curve we should recommend to an NBA general manager (GM). A GM's goal is to win the championship, so the curve he wants in an ideal world is something like the amount by which that drafting at position x increases the probability of winning the championship in the next t years. This is extremely difficult to estimate, so for now we want to choose the best proxy of this objective. We leave it to basektball teams to make that decision themselves.

2.5 Limitations

Our main contribution in this work is the creation of a novel Gamma regression model that maps NBA performance outcomes to salary, which allows us to compare various draft value curves by putting them on the same scale. Nonetheless, our analysis is not without limitations. In this work, we only explored three measures of basketball performance, BPM, RAPTOR, and WAR. A more comprehensive study would consider other outcome variables like VORP. We did not explore surplus value, or performance value minus cost, which accounts for the fact that later draft picks are paid much less than earlier ones. We explored just mean and median draft value curves, but it is possible that other objective functions may produce interesting and applicable draft curves. For instance, a team may be interested in a highly elite player, which necessitates using an objective function that values eliteness like right tail probability, or the probability a draft pick results in an elite player. Finally, the biggest limitation of our project is that we did not conduct a basketball discussion of which curves are right for which teams in which scenarios. We look forward to explorations of these enhancements in future work.

References

Baldwin, B. (2024). NFL Draft Value Chart: Constructing surplus value from scratch. https://opensourcefootball.com/posts/2023-02-23-nfl-draft-value-chart/.

Bossio, M. and Cuervo, E. (2015). Gamma regression models with the gammareg r package. *Comunicaciones en Estadística*, 8:211.

Fitzgerald, J. and Spielberger, B. (2024). Fitzgerald-Spielberger NFL Draft Trade Value Chart. https://overthecap.com/draft-trade-value-chart.

Massey, C. and Thaler, R. H. (2013). The loser's curse: Decision making and market efficiency in the national football league draft. *Management Science*, 59(7):1479–1495.

Pro Football Focus (2024). The PFF draft value chart. https://www.pff.com/news/draft-pff-draft-value-chart.

Stuart, C. (2012). Draft Value Chart. https://www.footballperspective.com/draft-value-chart/.