Is the NBA Combine an Effective Predictor of In Game Performance?

Note: *This document is a streamlined description of the methodology and high-level findings. For a more detailed view replete with visualizations, detailed analysis and code that can be used to replicate the findings please refer to the Jupyter Notebook.*

Overview: every year before the draft the NBA holds a "combine", an evaluation event where NBA prospects come to have their raw athletic abilities evaluated, physical measurements taken and shooting skills tested. For example: an individual athlete's height, weight, wingspan, vertical leap, speed and agility would be measured, in addition to their shooting accuracy from various parts of the court. The thinking is that the data gathered at the combine will not only help NBA teams better evaluate potential prospects, but lesser known athletes can "increase their draft stock" by demonstrating that they're better skilled and/or more athletic than was previously thought. In recent years the value of the combine has been questioned, some athletes have opted out and superstar players like Kevin Durant have stated the belief that it is a waste time.[1] It's worth noting that Kevin had a historically poor showing at the NBA combine (he wasn't to complete a single rep in the bench press and was ranked near the bottom), but was still a high draft pick and has built a hall of fame career that's still in progress.

This project attempts to assess the value of the NBA combine by determining whether or not there is any relationship between the combine measurements and in game performance. I.e. is Kevin an exception or the norm?

In this analysis we will primarily use the following two statistics:

- Win Shares (WS): a statistic that attempts to quantify an individual athlete's total value by quantifying how many wins they contributed to their team. The behind idea this and similar statistics (like player efficiency ratio) is to have a single number to rank players with that compensates for different styles of play. E.g. some players score a lot of points but don't get a lot of rebounds, while others score fewer points but have higher rebound and assist numbers.
- Points per game (PPG)

The research questions for this project will be as follows:

1) How good are the NBA's methods for evaluating talent overall? Meaning: is there a strong relationship between draft position and in game performance? Are *"draft busts"* where players significantly underperform relative to their draft position rare or common?
2) Is there a significant relationship between draft position and combine performance? Going to Kevin Durant's statements, does the combine hold any value for the athletes?

---

[1] ABC News (ABC News Network), accessed August 9, 2020, https://abcnews.go.com/Sports/kevin-durant-calls-nba-combine-waste-time-top/story?id=47338234.

By. Markham Lee – September 2020

3) Are there any relationships, patterns or strong correlations between NBA combine performance and in game statistics like PPG and WS? E.g. do faster and more agile players score more points?

In addition to the above, we will look at other statistics if we discover they do in fact have a strong relationship with combine measurements/performances, E.g. wingspan and rebounds.

Caveats and Assumptions

1) This is very much a "prototype" analysis where one looks at a subset of data to better inform a more comprehensive analysis of a larger dataset. Potential future iterations could potentially look at 20-30 years' worth of NBA statistics and combine measurements.
2) Not all NBA players attend the draft and when they do not all of them participate in every event or measurement, which means that the analysis will be looking at a subset of the athletes that played in the NBA or a given time period.
3) We only included player seasons where they played in at least 58 games as that's the number of games required to have a "statistically significant season", in terms of the athlete qualifying for statistical rankings and awards like the scoring title.
4) The combine analysis only included data for 99 players, however, in a given point in time there are typically only 360 active NBA players, so 99 players is a reasonable sample size.
5) No adjustments were made for situations where injuries may have severely impacted a player's numbers even though they played in 58 games, nor did we look at longevity or scenarios where a player had some great seasons and then some poor ones, so their career averages are mediocre. However, this may be addressed in future versions of this project.
6) When evaluating draft position vs. game performance I only looked at players that I had combine data for, i.e. it doesn't include all players drafted since not all of them participated in the combine.
7) A correlation coefficient is considered strong if it's +/- 0.70

High Level Findings ***(detailed findings, visualizations and the like are in the Jupyter Notebook)***

1) There didn't appear to be a strong relationship between any of the combine measurements and win shares or PPG, as no a single combine measurement had a correlation coefficient with WS and PPG that exceeded the absolute value of 0.26
2) The +/- 0.70 threshold wasn't reached for a single combine measurement any NBA in game statistic. The closest to 0.70 were for wingspan and blocks (0.67), wingspan and total rebounds (0.59) and height and blocks (0.57).
3) The highest performing players are nearly always high draft picks, but the chances of a high draft being above average is at best 50/50.
4) Only 33% of the athletes taken in the top 15 draft picks were above average in terms of win share
5) 43/93 athletes taken in the top 15 of the draft were above average in terms of PPG, however, a statistical significance test revealed there wasn't a meaningful difference

between the proportion of athletes above or below average scoring wise. Meaning: whether or not a particular athlete is going to be an above average scorer is basically a coin flip.

6) Only 19.35% of the athletes taken in the top 15 of the draft were in the 75th percentile for win shares

7) Only 27.96% of the athletes taken in the top 15 of the draft were in the 75th percentile for PPG

Due to the low correlations and the lack of a strong trend in the linear model plots, neither machine learning nor regression was pursued. While it's possible that a model could be created that considers a non-intuitive or not readily apparent relationship between certain metrics, the math doesn't support a strong relationship to build a model on, at least not without the inclusion of additional variables.

Methodology:

In order to not bias the analysis towards a particular playing style, I selected one of the composite metrics that are used to make *"apples to apples"* comparisons between athletes with different playing styles. I selected win share (WS), a metric that attempts to quantify an athlete's contribution in terms of number of wins, because it seemed less vulnerable to anomalies and manipulations than other composite metrics. For example, the player efficiency ratio (PER) rewards players for missed shots:

> *"A three-point field goal made is worth 2.65 points. A missed field goal, though, costs a team 0.72 points. Given these values, with a bit of math we can show that a player will break even on his two-point field-goal attempts if he hits on 30.4% of these shots. On three-pointers the break-even point is 21.4%. If a player exceeds these thresholds, and virtually every NBA player does so with respect to two-point shots, the more he shoots the higher his value in PERs. So a player can be an inefficient scorer and simply inflate his value by taking a large number of shots."(Bleacher Report - https://bleacherreport.com/articles/1040320-understanding-the-nba-explaining-advanced-comprehensive-stats-and-metrics)*

In comparison win shares is a fairly strong metric that has few drawbacks beyond not considering playing time per game, which can be easily adjusted for by expressing the number as win share per 48 minutes (length of an NBA game).

The analytical approach was as follows:

1) Use NBA player statistics data from the 2009 – 2010 seasons through the 2018 – 2019 seasons
2) Drop statistics that are strongly correlated or are functions of other metrics, E.g. Field Goal %, total rebounds as opposed to offensive and defensive rebounds and

3) In the NBA dataset, filter out seasons where the athlete didn't play at least 58 games, as 58 games is the threshold for a player to have a *"statistically significant"* season as far as rankings, scoring titles and the like.
4) Roll up the NBA dataset by player and average statistical performance, so we have the average points scored, rebounds, win shares, etc., over the time period
5) Use combine data from the 2009 through 2017 combines, this will ensure that even the "newest players" will have at least two NBA seasons
6) Drop statistics that are either somewhat redundant or highly correlated. E.g. there were multiple measurements types for reach and vertical leap, height and wingspan were highly correlated.
7) Merge the datasets by player so I have a combined dataset with the athlete's average performance per season and their combine numbers.
8) Create correlation matrixes to gain an understanding of the relationship between various metrics (E.g. draft position vs. win shares) and then use that to identify variables for linear model plots and potentially linear regression and/or machine learning models.
9) Calculate the mean PPG and WS and use that to evaluate what % of high draft picks (defined as being chosen in the top 15 of all players in a given draft) perform above average.
10) Repeat the analysis above, only using the 75th percentile of WS and PPG to see what % of high draft picks perform above that threshold.

Data Sets:

1) NBA draft combine dataset: https://data.world/achou/nba-draft-combine-measurements
2) Basketball Reference: https://www.basketball-reference.com/

Testing

- Since the project was composed of visualizations, correlation coefficients and linear model plots it didn't lend well to traditional test cases. I validated that my joins and the like were correct by visual inspection of the NBA player data with a similar data from Basketball Reference. I.e. if my top 30 players for win share matched Basketball Reference top 30 for the same time period, then the data was imported and appended into a single data frame correctly.
- For the combine joins I visually inspected my combined data set with the athlete's combine performance from the original data set, plus their average statistics from Basketball Reference. E.g. Damian Lillard's combine performance in my data set matched the original perfectly and his avg PPT was within 0.02 between my Pandas Data Frame and Basketball Reference, given that the avg PPT in my dataset is an average of an average and the number in basketball reference is taking total points and dividing them by total gains, the difference was minor given the different approaches to calculating PPG.

Data Dictionary

Used the following data fields from the NBA player data: **ORB and DRB were in the data set but weren't used*

Player
Season
Win shares (WS)
Games (G)
Minutes Played (MP)
Offensive Rebounds (ORB)
Defensive Rebounds (DRB)
Total Rebounds (TRB)
Assists (AST)
Steals (STL)
Block (BLK)
Points (PTS)
Turnover (TOV)
Points per Game (PPG)

Most of the other measures were "rates" as far as free through %, field goal percentage, average minutes per game and the like, all of these measures would have high correlations with the measures I kept so it didn't make sense to use them.

Combine Data Measures Included ****fields highlighted in green were the ones that were used in the analysis***

Player
Draft Year
Draft Pick
Height (No Shoes)
Height (With Shoes)
Wingspan
Standing reach
Vertical (Max)
Vertical (Max Reach)
Vertical (No Step)
Vertical (No Step Reach)
Weight
Body Fat
Hand (Length)
Hand (Width)
Bench
Agility
Sprint

By. Markham Lee – September 2020

The fields not used were rejected due to either being a variant of the same thing (height without shoes vs. height with shoes), repetitive or a function of another variable: e.g. Vertical (Max Reach) vs. Vertical (Max). There were other measures that were rejected due to the players turning in similar performance. For example spot up shooting where the players each shoot five balls from a particular spot on the floor: few athletes missed more than two shots and most missed one or fewer, so there wasn't enough differentiation between the individual athlete's performances to justify using it.

By. Markham Lee – September 2020