Using NBA Draft Combine Results to Predict Player Success

Kyle Riebeling

2025-03-17

Abstract

This analysis compares players NBA Draft Combine stats to their NBA career stats to find any significant predictors for better performance. I used NBA Draft Combine stats from 2001-2019 and their matching career stats. In this analysis, we split players into subgroups based on their primary positions. Using linear regression modeling, we find that guards with better sprint times have increased win shares, wings with higher bench press reps have lower win shares, and bigs with higher jumping reach have increased win shares. All other correlations were insignificant. This shows that different metrics are have differing importance between the subgroups. We can also conclude that athletic metrics, like spint times, are more important in predicting win shares compared to body measurements, like wingspan and height.

Introduction

Each year the NBA hosts a Draft Combine. The Draft Combine is an event where prospective NBA players are invited to show their skills. It involves measuring their physical attributes, like height and wingspan, and testing their physical abilities, like timing their reactive shuttle run and measuring their vertical leap. Drafting the right player can turn a the future of a franchise around. It can be hard to find these franchise players, however. Often we see players drafted early end up being mediocre or flat out horrible in the NBA. We also often see generational MVP-level players drafted quite late into the draft, like this year's MVP candidates Giannis Antetokounmpo at 15th and Nikola Jokić at 41st during a Taco Bell commercial.

The difficulty of drafting has increased as the talent level of all NBA players has increased in recent years. It has been over 10 years since the last 1st pick in the draft won MVP, when Lebron James won it in 2013. This shows how common it is to skip someone with MVP potential in the draft and the importance of analyzing their combine metrics to ensure you don't skip someone with high potential. My goal in this analysis is to compare players combine statistics to their success in the NBA. I will run the same analysis on three subsets of positions to see if there are different measurements that matter more for different types of players

Data Used

For this analysis, I am using two data sets from Kaggle. The first data set is a collection of NBA Draft combine stats starting from the year 2000 (https://www.kaggle.com/datasets/marcusfern/nba-draft-combine). It includes anthropometric, strength, and agility statistics. sometimes players that go undrafted one year will be invited back. In this case, only the most recent measurements are kept.

The second data set was a bit more tricky to pick. I needed a data set that would measure an NBA player's career success. Often, simple statistics, like points per game and box plus-minus, are not capable of truly showing a players value to his team. Several statisticians have attempted to create advanced statistics that are calculated using a combination of a players simple statistics to better represent their value. There are many of these advanced statistics to choose from, but I decided to choose "win shares". Win shares attempt to show how much a player contributes to each of their team's wins. To see how this value is calculated,you can visit Basketball Reference's description page (https://www.basketball-reference.com/about/ws.html). Essentially, one team win is equivalent to one win share - the total win shares a player has is the amount of games he won for his teams throughout his career.

Once I decided to use win shares over other advanced statistics, I found another data set that included player names and win shares on Kaggle in order to match it with my combine statistics data set. I then changed the naming schemes to match and merged them into one data set.

Subsetting

Once I had my combined data set, I created three subsets based on the players primary and secondary positions. The NBA is becoming more of a position-less league, so using the traditional five positions would cause this analysis to be outdated. I decided to split the players into three general positions - wings and guards. I defined wings as players who had either their primary or secondary position listed as small forward (SF). Bigs are players who have the primary position power forward (PF) or center (C). Lastly, all other players were placed into the guards category and had either point guard (PG) and shooting guard (SG) as their primary positions.My final subsets of data for the analysis include 182 bigs, 166 wings, and 163 guards with no missing entries.

Combine Metrics Selected

The NBA draft combine measures many things, too many for the purpose of this analysis. Some measurements are also often missing each year. In order to perform a useful analysis, we need to select a few measurements in this data set that don't have many missing entries and provide information on multiple aspects of each player. I selected the following four:

- sprint: This variable measures the time it takes a player to sprint 3/4th of the court in seconds
- bench: This variable measures the amount of 185lbs bench press repetitions a player was able to do
- bar: This variable is the ratio of a players wingspan to height
- pdhgt: This variable measures the highest a player can reach when they jump in inches

These four variables offer insight into a players speed, jumping ability, strength, and length - all valuable things to have as a basketball player. While there are other variables in this data set that measure other useful things as well, there were too many missing entries. Using other metrics would cause our sample sizes to decrease greatly.

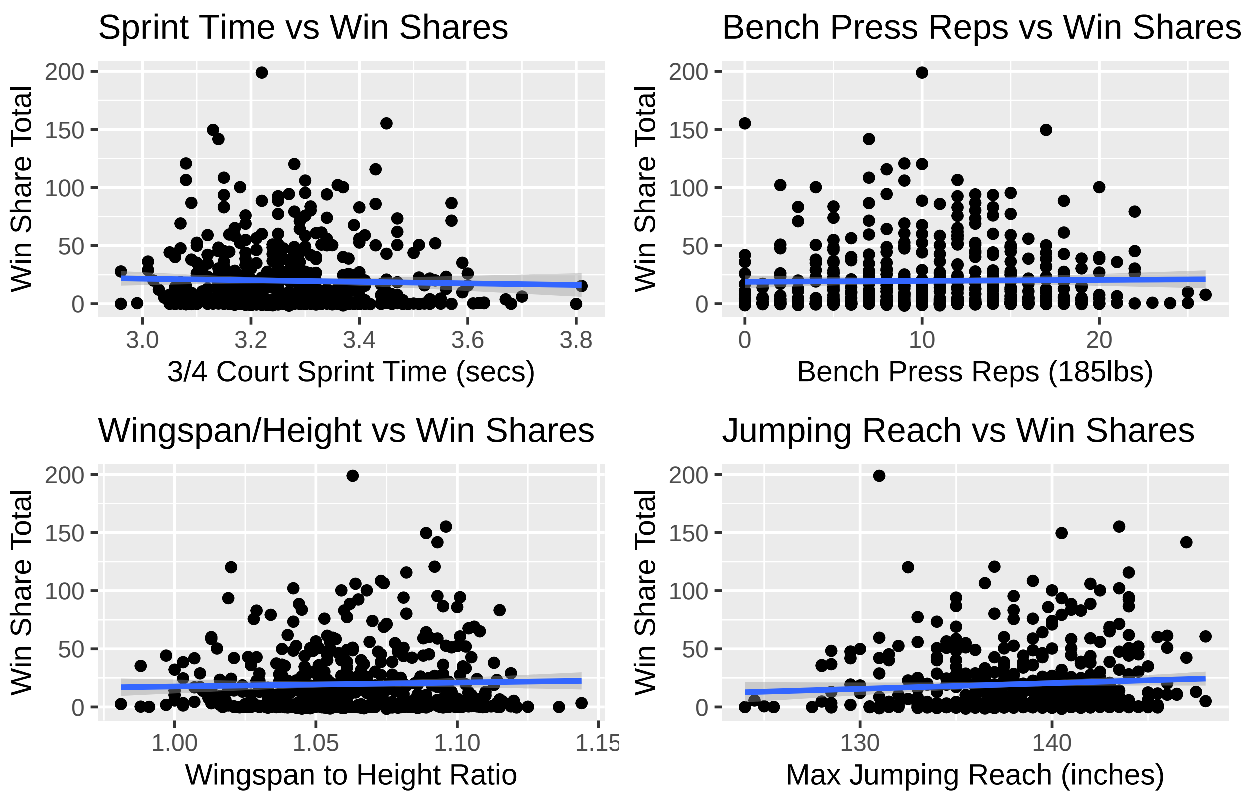
Analysis Method

To answer our question, we run a linear regression model for each of the 4 variables where y is win shares and x is one of the variables. We then plot these results side by side to see these relationships. Any model that finds significant results will also display the p-value and coefficient in the top right of the graph to help distinguish between significant and insignificant results.

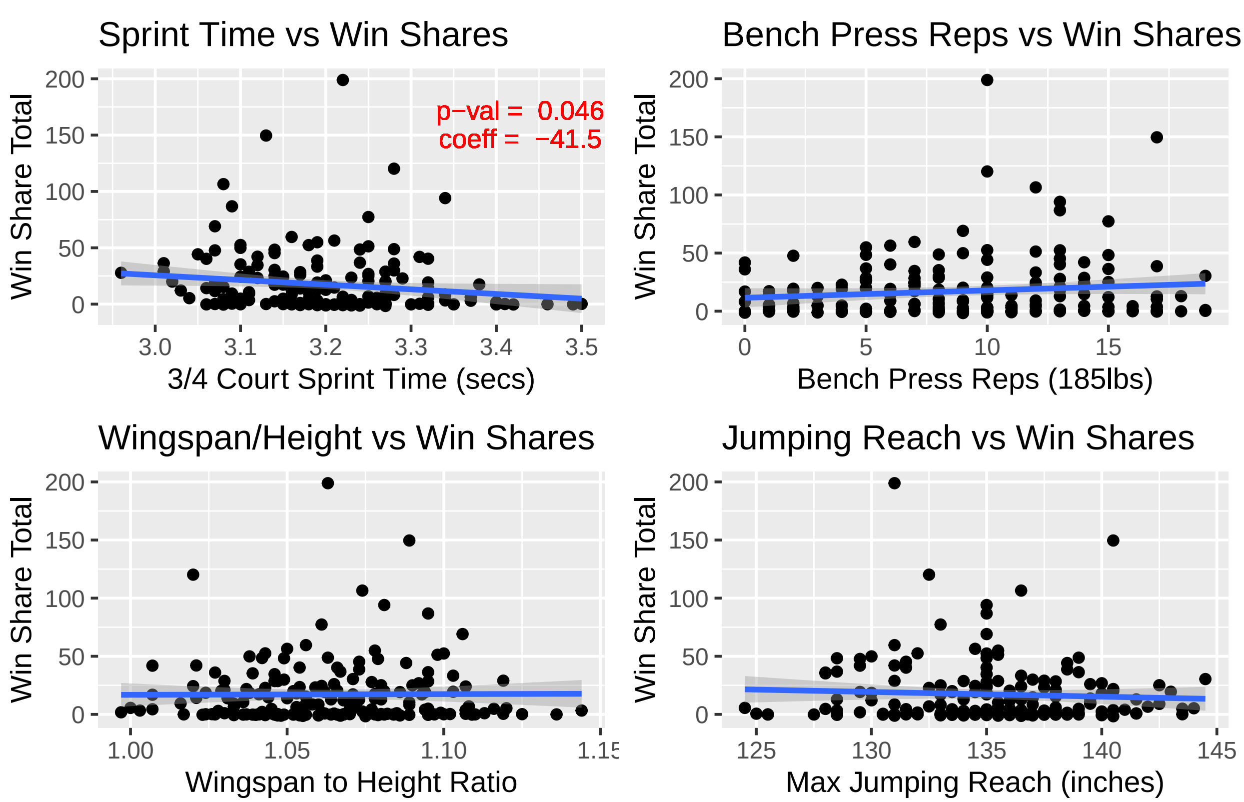
This process is done for all players in our data set and then repeated for our 3 position subgroups to highlight any differences between them.

Visualizaion

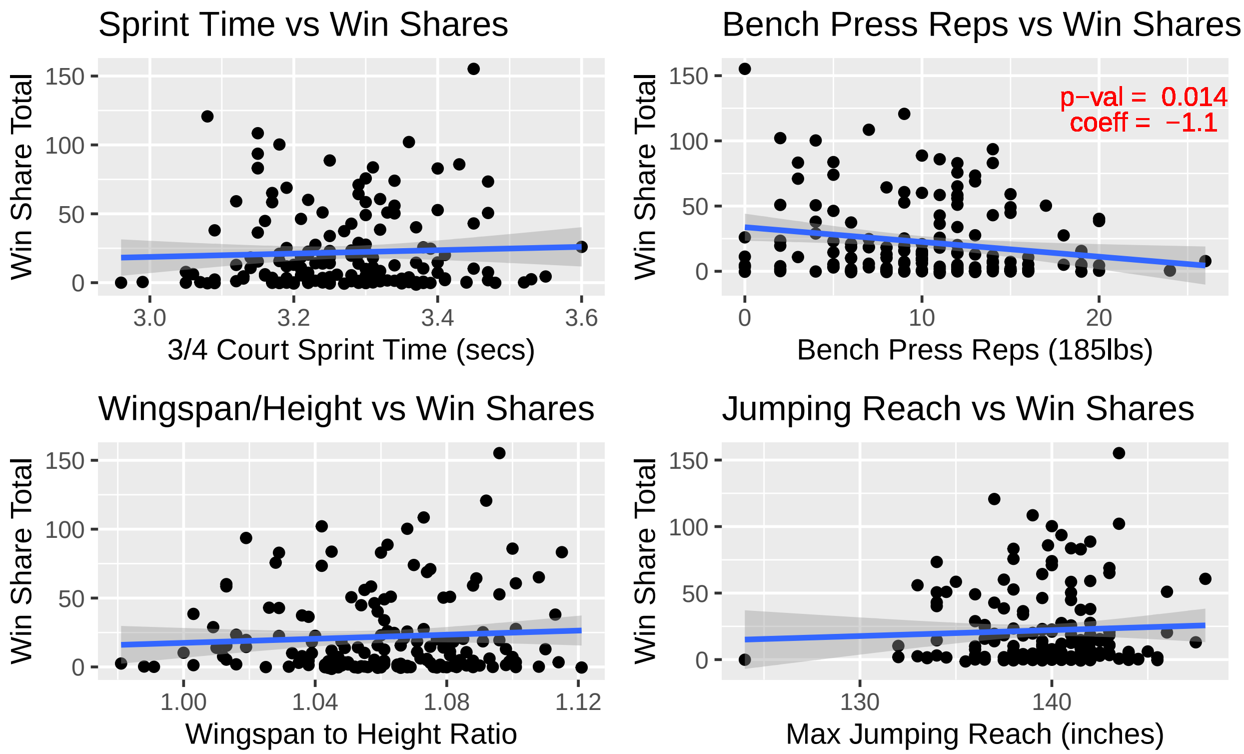
All Players


Sprint Time vs Win Shares
Win Share Total
3/4 Court Sprint Time (secs)
Bench Press Reps vs Win Shares
Win Share Total
Bench Press Reps (185lbs)
Wingspan/Height vs Win Shares
Win Share Total
Wingspan to Height Ratio
Jumping Reach vs Win Shares
Win Share Total
Max Jumping Reach (inches)

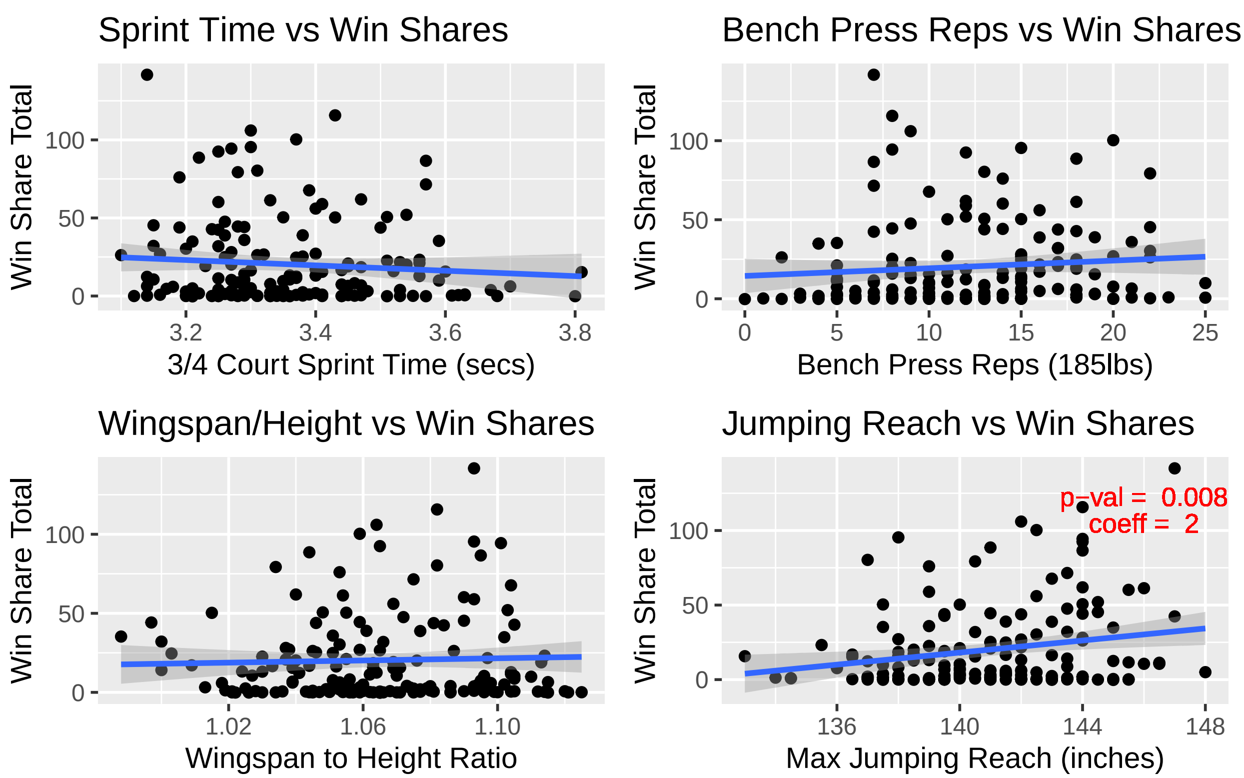
Guards


Sprint Time vs Win Shares
p−val = 0.046
coeff = −41.5
Win Share Total
3/4 Court Sprint Time (secs)
Bench Press Reps vs Win Shares
Win Share Total
Bench Press Reps (185lbs)
Wingspan/Height vs Win Shares
Win Share Total
Wingspan to Height Ratio
Jumping Reach vs Win Shares
Win Share Total
Max Jumping Reach (inches)

Wings


Sprint Time vs Win Shares
Win Share Total
150
100
50
0
3.0
3.2
3.4
3.6
3/4 Court Sprint Time (secs)
Bench Press Reps vs Win Shares
p−val = 0.014
coeff = −1.1
Win Share Total
150
100
50
0
0
10
20
Bench Press Reps (185lbs)
Wingspan/Height vs Win Shares
Win Share Total
150
100
50
0
1.00
1.04
1.08
1.12
Wingspan to Height Ratio
Jumping Reach vs Win Shares
Win Share Total
150
100
50
0
130
140
Max Jumping Reach (inches)

Bigs


Sprint Time vs Win Shares
Win Share Total
100
50
0
3.2
3.4
3.6
3.8
3/4 Court Sprint Time (secs)
Bench Press Reps vs Win Shares
Win Share Total
100
50
0
0
5
10
15
20
25
Bench Press Reps (185lbs)
Wingspan/Height vs Win Shares
Win Share Total
100
50
0
1.02
1.06
1.10
Wingspan to Height Ratio
Jumping Reach vs Win Shares
p−val = 0.008
coeff = 2
Win Share Total
100
50
0
136
140
144
148
Max Jumping Reach (inches)

Results

When running the analysis for all players in our data, we get no significant results and the regression lines are almost completely flat. Each of our subsets have significant results in different categories, however. This proves that it is valuable to do this analysis separately for each type of player.

Guards

For guard players, bench press repetitions, wingspan to height ratio, and max jumping reach were all insignificant in predicting win share totals. 3/4th court sprint times, however, was found to be significant with a p-value of 0.046 and a coefficient of -41.5. This makes sense as a guard's role in basketball is to move around the court and facilitate offense for other players. Everything important to what a guard does is easier to do when they are faster than other players. Since guards are typically shorter and weaker than other players on the court, they need to rely heavily on their speed to get open shots of create space for other players to score. Having the speed to blow by bigger and slower defenders is a skill many great point guards posses and take advantage of regularly.

Wings

For wing players, 3/4th court sprint time, wingspan to height ratio, and max jumping reach were all insignificant in predicting win share totals. This time, bench press repetitions was the only significant predictor with a p-value of 0.014 and a coefficient of -1.1. This is a bit perplexing. It suggest that being stronger than other players is actually a negative thing for wing players. This might be because higher bench press repetitions means higher muscle mass and decreased agility. Further analysis into this outcome might offer a better explanation. Surprisingly, 3/4th court sprint times were not significant for wing players. While they are taller typically taller and stronger than guard players, they are still out sized by bigs. Wings are often more defensive focused players, so it is possible that we would get more meaningful results for wings if we ran these models against a defensive focused advanced statistic rather than win shares.

Bigs

For big players, 3/4th court sprint time, wingspan to height ratio, and bench press repetitions were all insignificant in predicting win share totals. Max jumping reach was significant in predicting win shares with a p-value of 0.008 and a coefficient of 2. This result makes sense as a big's main role is to grab rebounds and be a defensive anchor on shots taken close to the basket. Having a higher jumping reach helps with both of these skills.

Bench press repetitions not being significant for big men was a bit surprising to me. Bigs take advantage of both aspects of their physical size when playing - height and strength. They often post us and push back smaller defenders that aren't strong enough to hold them back in order to get an easier shot near the basket. This could be a reflection of the change to a position-less type of play, however - giving up pure strength to have increased speed is an increasing trend for big men in the NBA. For example, Victor Wembanyama is one of the leauge's new big men with the most potential. He has the nickname "Slender Man" because his stature resembles the tall and lanky character. While he does get overpowered often, his slender frame allows him to play similar to a guard but with the height of a big man. This new type of big man is becoming more common and might be the reason bench press repetitions were found not significant for bigs.

Conclusion

Splitting the players into subgroups proved to be valuable in this analysis. Each group had a different combine metric that was significant in predicting win shares. A summary of these findings are as follows:

- Guards: The slower the sprint time, the fewer predicted win shares
- Wings: The more bench press repetitions, the predicted fewer win shares
- Bigs: The higher jumping reach, the more predicted win shares

For the guard and big subgroups, the findings matched what would be expected. For wings, however, the findings were perplexing. We also found no significant correllation between all subgroups and wingspan to height ratio. This suggest that athletic metrics are more valueable that physical measurments. This makes sense as physical measurments are just "tools" while athletic metrics show how well the players can utilize their "tools".

Areas for Further Analysis

It would be helpful to look into the negative association between bench press repetions and win shares in wings. It is possible that there is another variable measured during the combine that is related to bench press that we excluded from this analysis that would better explain why higher strength leads to decreased win shares. Bench press repetitions may be associated to lower shooting percentages or lower agility times which would better explain why being stronger would predict lower win shares

Analysing with a different advanced statistic that focuses more on defense would also likely be more helpful for the wing subgroup, as wing players are often more valuable on the defensive side of basketball due to their versitile size. The "3 and D" archetype is common for wing players - a player whos offensive value is simply being a good catch and shoot 3 point player, while most of their value comes from being able to play defense on players of all sizes.

Another interesting finding was that bench press is not significant for big players in predicting win shares. I believe this is due to the shift in prioritizing skinnier, more agile big men with guard skills over stronger, slower big men that depend on their size. I believe running this same analysis for big men from an earlier period in NBA history and comparing them to the results of the big men in this anaylsis would provide much different results. The shift to a "position-less" NBA has mainly effected the big men in the league, so it is likely that these same metrics hold different values now compared to an earlier period.