The Pascal Siakam Model

Finding Diamonds in the Rough in the NBA Draft

By Ayush Batra, Wesley Neidhardt, Pranav Nanda, and Daniel Galper

Wharton Moneyball Academy 2020, Sports Analytics Student Research Journal

The Problem

- NBA Draft picks performances in the NBA can almost feel random at times

- Why do non-lottery picks like Pascal Siakam and Kawhi Leonard go on to have All-Star careers while top draft picks like Kwame Brown, Derrick Williams, and Anthony Bennett end up as busts?

Our Goal

- Create a model that uses NBA players standard and advanced stats from college to understand why they were successful or unsuccessful in the NBA

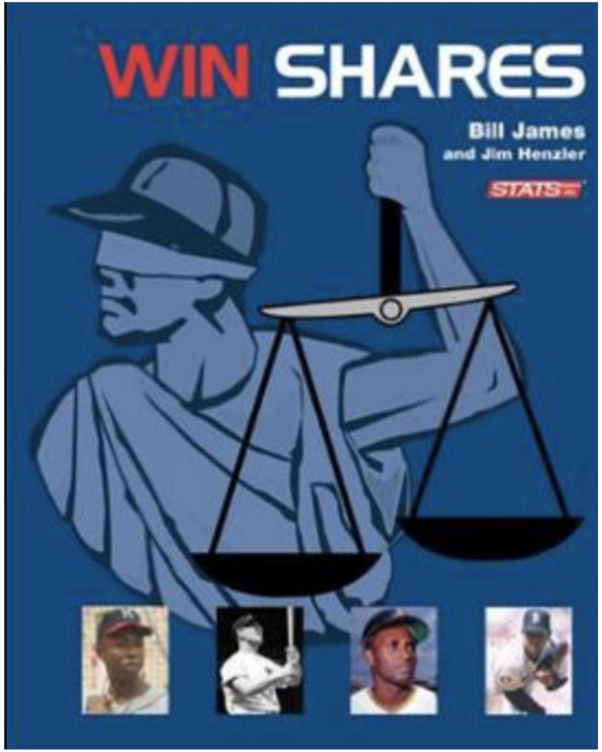
- Apply that model to current college players and try to predict their success in the NBA

Methods

- Gather college and NBA data from Basketball-Reference for every player drafted in the 2007-2016 drafts who played in college and in over 100 games in the NBA

- Use R to fit and build regression models that predict standard stats like 3-point percentage and advanced measurements like Win Shares.

Win Shares



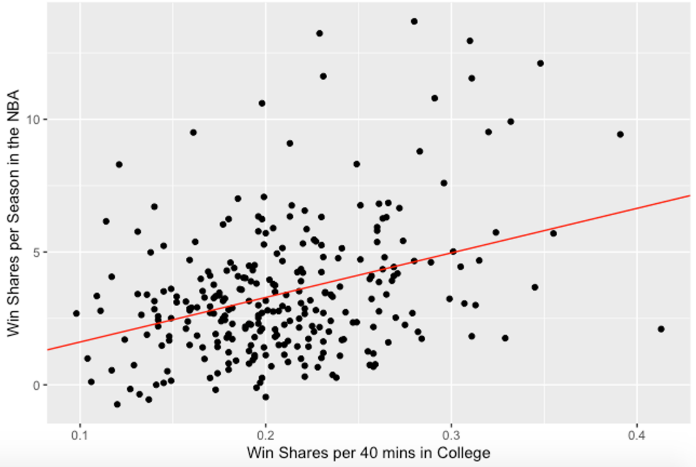
- A metric that attempts to assign an amount of wins a player contributed to their team
- Main measurement we used to calculate a players success in the NBA
- We used Win Shares/40 minutes for college stats and Win Shares/82 games for NBA stats to measure how many wins each player added over a full 82-game NBA season

Initial Models

First steps towards finding which college stats matter the most when predicting NBA stats

Predicting NBA Win Shares

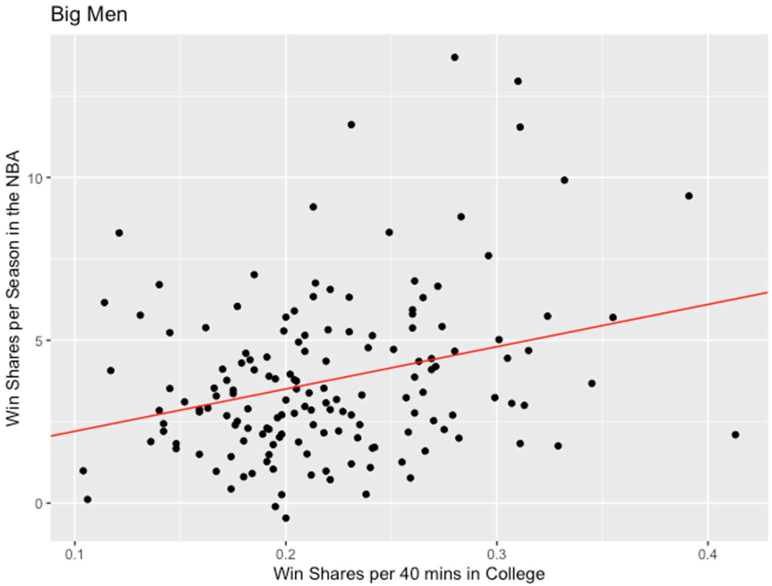
- First we converted win shares in the NBA to win shares per season
- We then used win shares per 40 minutes in college to predict win shares per season in NBA
- Although college win shares was significant in predicting win shares in the NBA, the r-squared value was fairly low



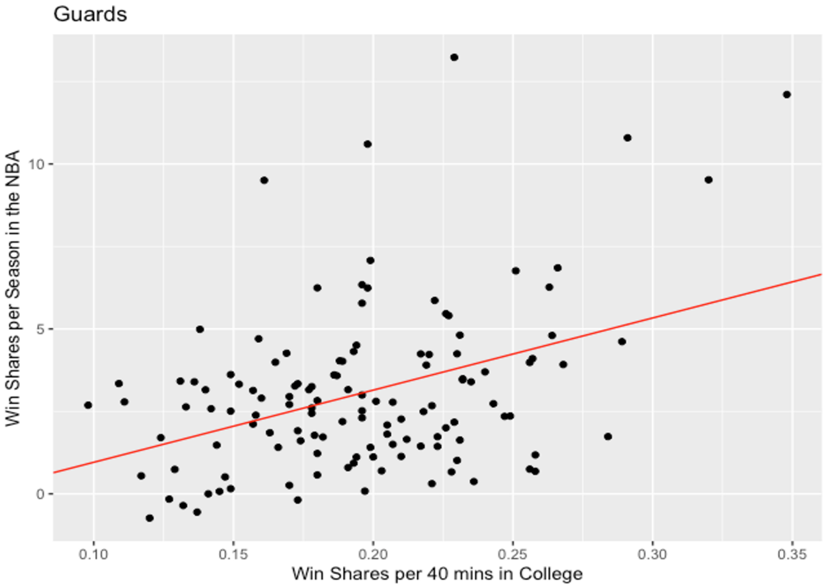
R-squared = 0.1249
Correlation = 0.3534

Predicting NBA Win Shares

- We then split the players into 2 groups: guards and 'big-men'



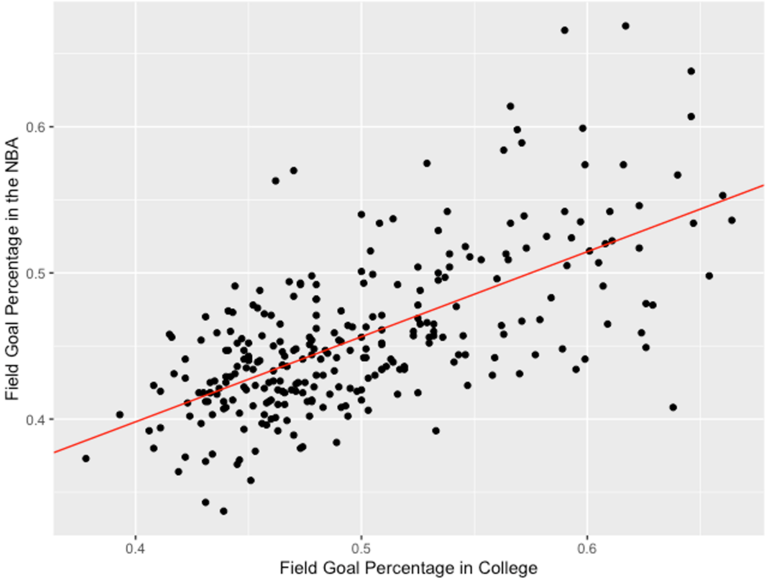
R-squared = 0.08553
Correlation = 0.2925



R-squared = 0.1593
Correlation = 0.3991

Comparing FG%

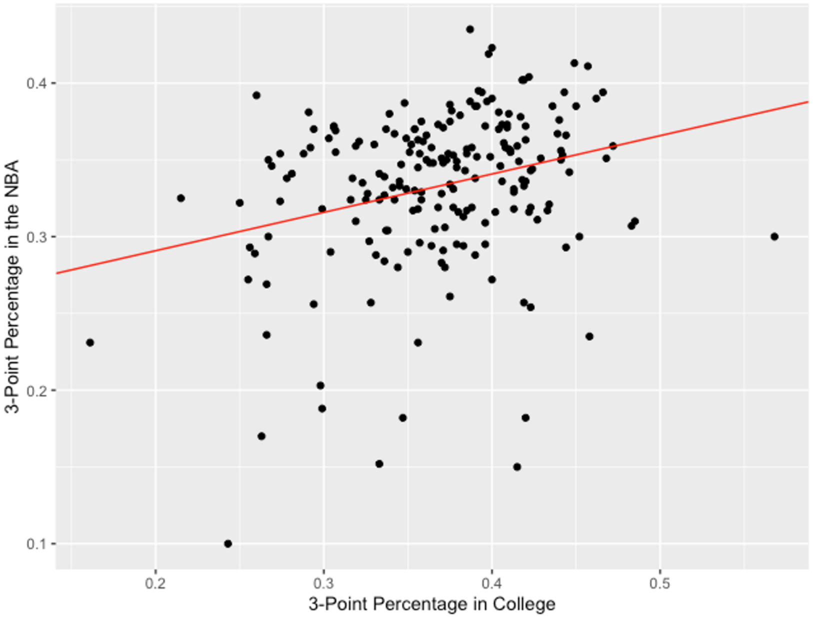
- We also compared 3 NBA shooting stats to the same college stats: FG%, 3P% and FT%



R-squared = 0.4225
Correlation = 0.650

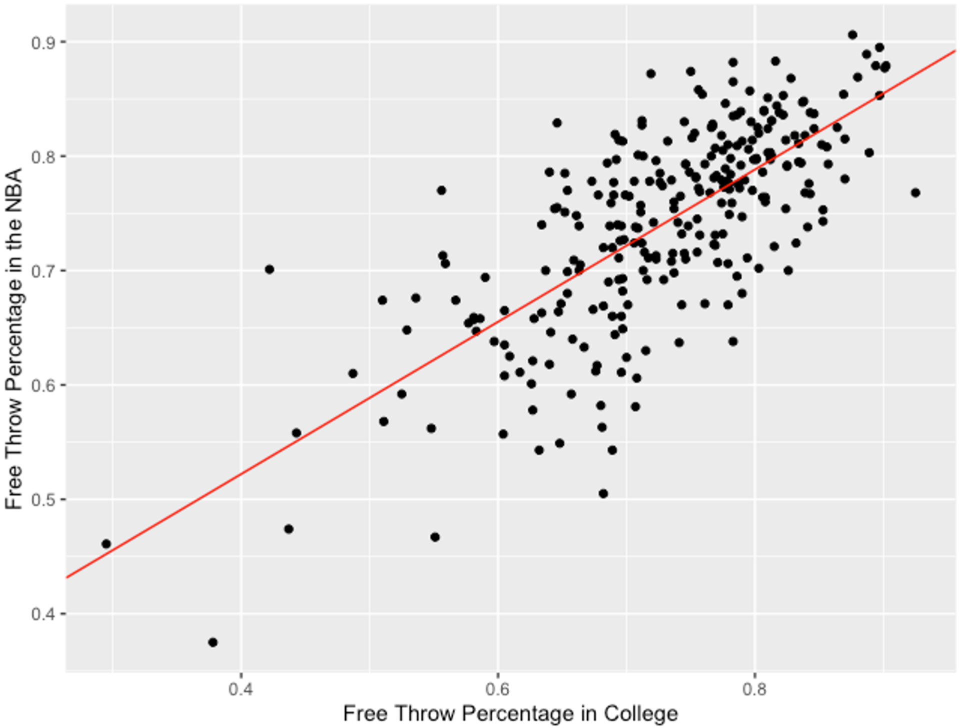
Comparing 3P%

- We filtered the dataset so only players who attempted more than one 3-pointer per game in college were included



R-squared = 0.07727
Correlation = 0.2780

Comparing FT%



R-squared = 0.5270
Correlation = 0.7259

Conclusions and Next Steps

- Splitting the players into groups based on their positions will give us a better prediction for NBA win shares
- College FT% and FG% are much more significant in predicting NBA FT% and FG% than college 3P% is in predicting NBA 3P%
- A multiple regression model is likely to be more effective than a univariate regression model

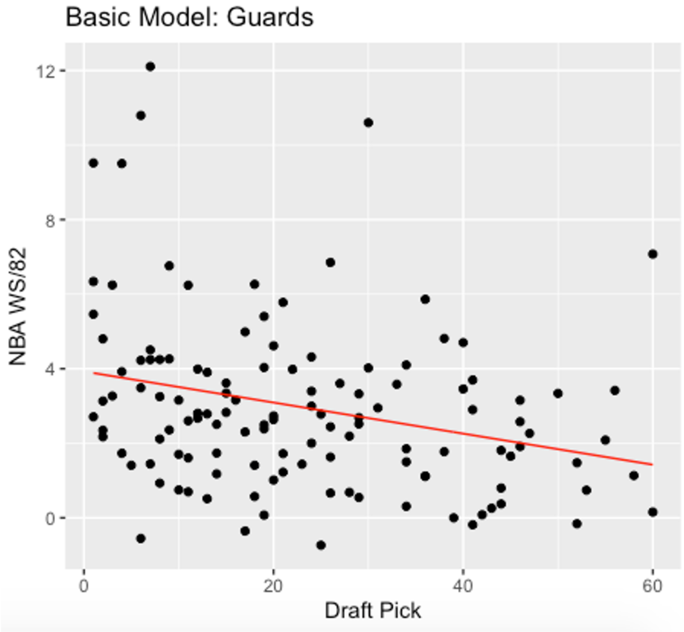
Final Models

Predicting NBA Win Shares and Shooting Stats

Working With Guards

- First, we split the dataset of players into two groups: Guards and Big Men
  - Unfortunately, there were an insufficient amount of Centers for their own group

- Then, we looked at the relationship between Draft Pick and Win Shares per 82 games



R-sqrd = 0.078

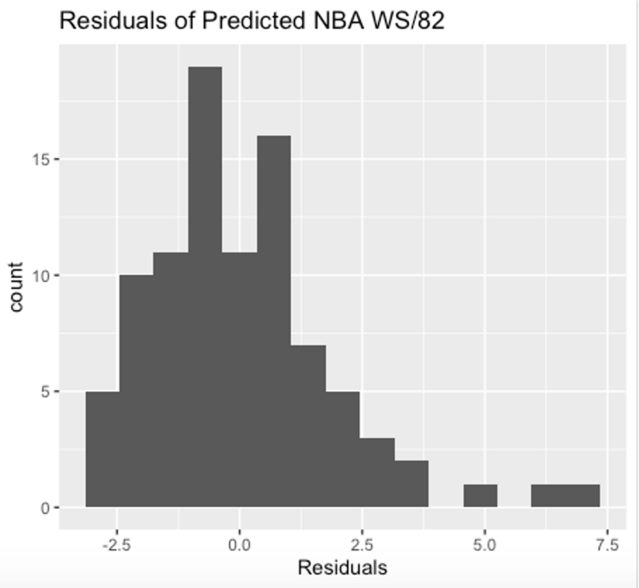
WS Regression: Guards

- We created a multiple linear regression using college stats as inputs in order to predict the players’ NBA WS/82

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -14.20691    4.03014  -3.525 0.000684 ***
Pk           -0.03230    0.01293  -2.499 0.014376 *
`3P`         -5.74160    1.88960  -3.039 0.003158 **
`3PA`         2.20698    0.73780   2.991 0.003634 **
TSP          17.92503    7.20916   2.486 0.014860 *
`3PP`        16.20023    7.53424   2.150 0.034379 *
FTr           3.23173    1.98357   1.629 0.106961
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.926 on 85 degrees of freedom
Multiple R-squared:  0.2677,    Adjusted R-squared:  0.216
F-statistic: 5.178 on 6 and 85 DF,  p-value: 0.0001397
```

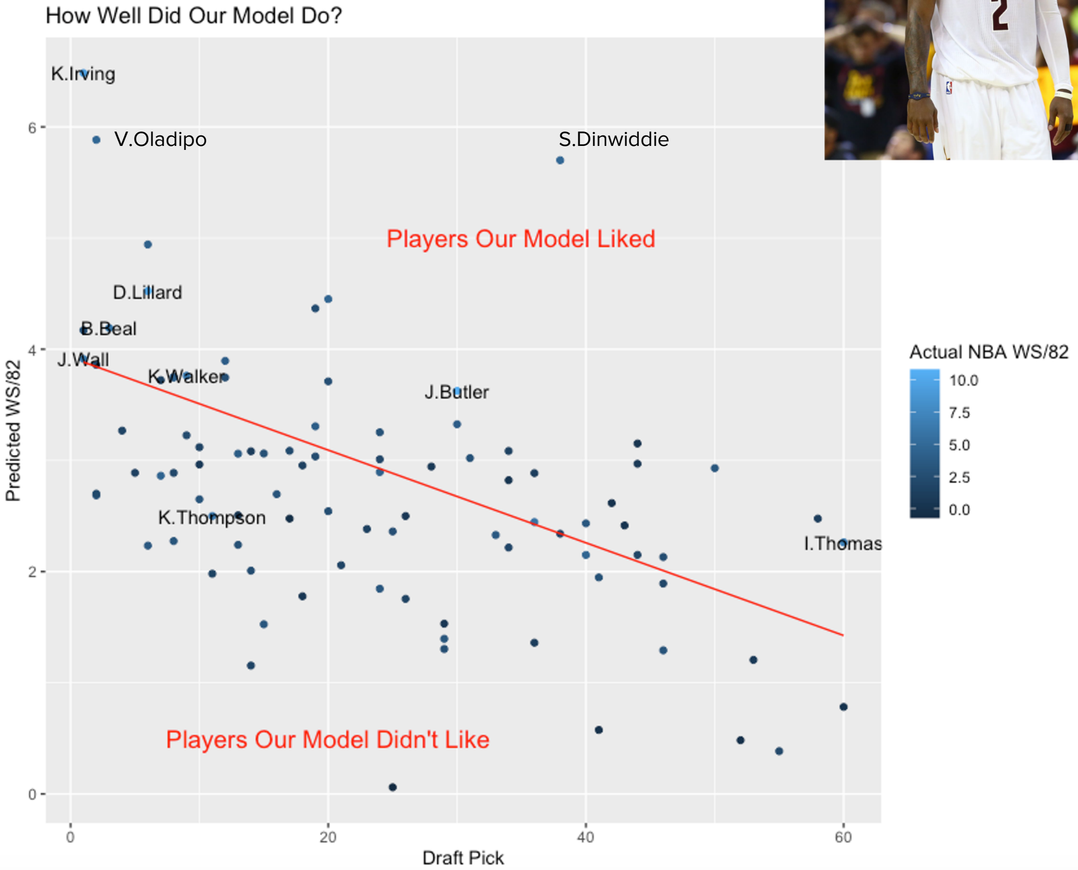
Results of Multiple Linear Regression



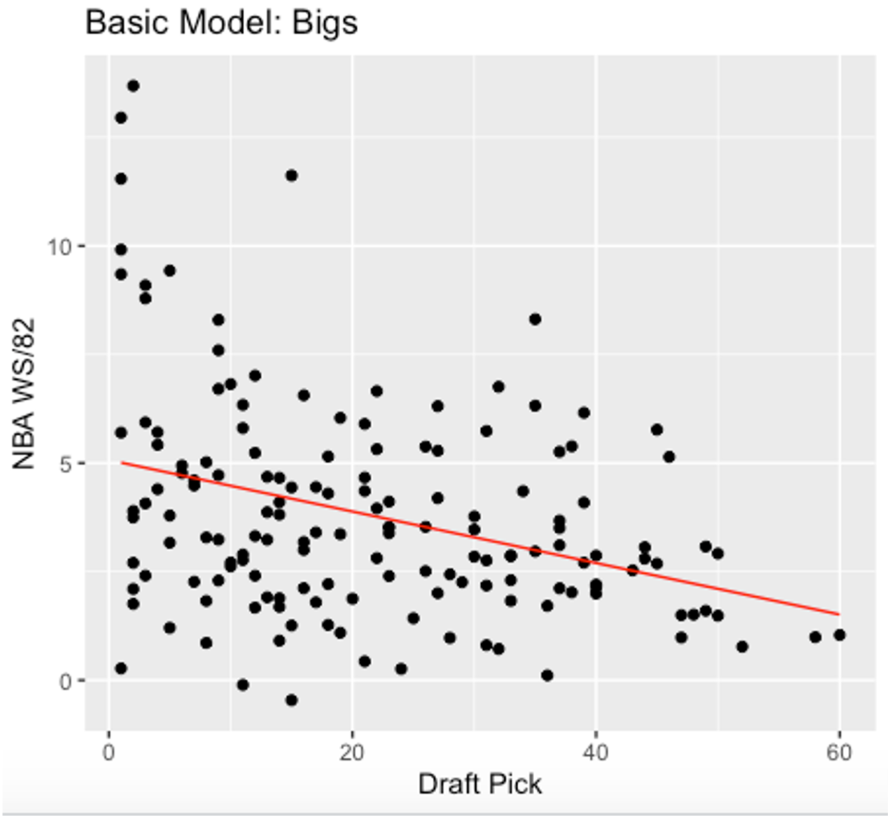
Unfortunately, the residuals are not normal
Bad Model?

Model Performance: Guards

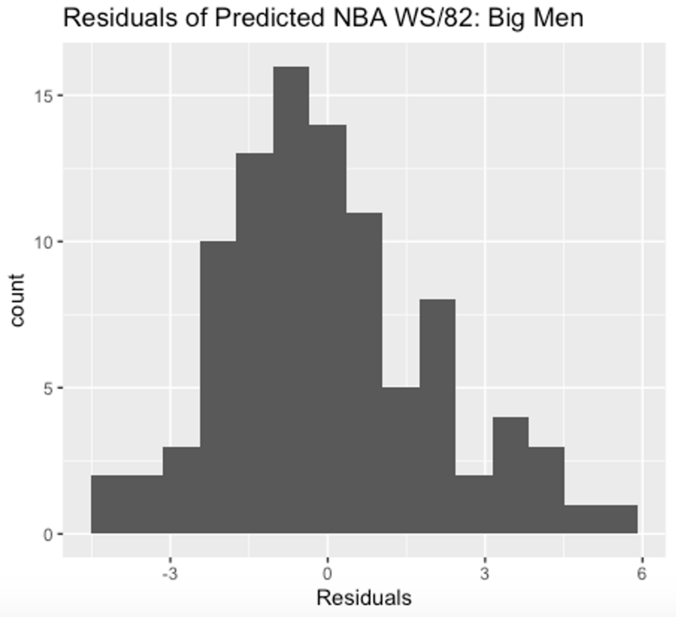
- Let's see how our model performed on past draft prospects
- Players above the line are those who the model expected to outperform their draft position (and vice versa)



WS Regression: Big Men



Residuals = Approx. Normal

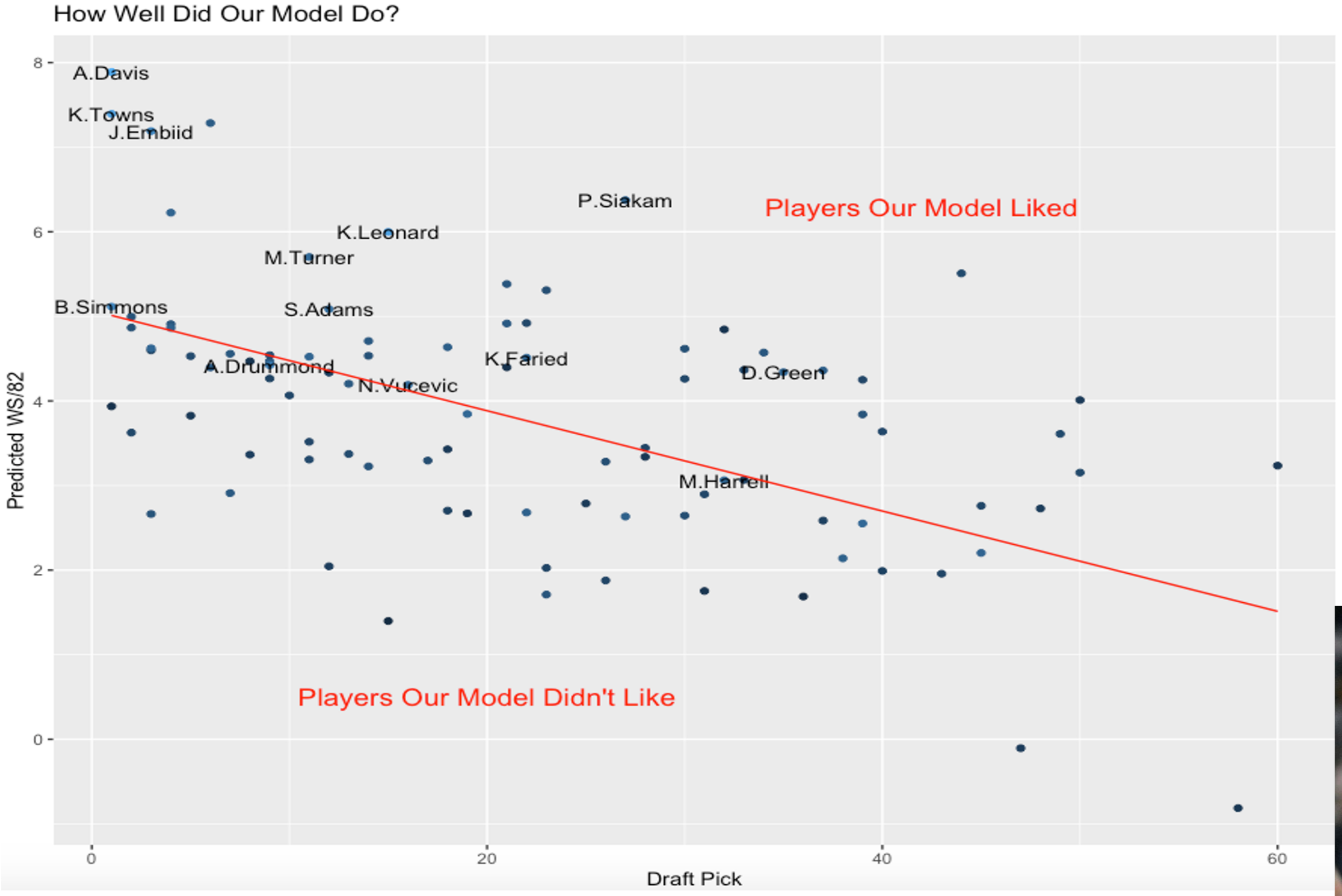



```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.40299    3.24351   1.974 0.051544 .
College_Yrs   -0.66879    0.19189  -3.485 0.000773 ***
FTA           -1.95534    0.70488  -2.774 0.006775 **
TSP          -24.26228    7.63948  -3.176 0.002067 **
FTr           15.30918    5.10880   2.997 0.003557 **
PProd_G        0.76215    0.25337   3.008 0.003437 **
BLKP           0.20231    0.08298   2.438 0.016795 *
`WS/40`       17.83026    6.72158   2.653 0.009490 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.086 on 87 degrees of freedom
Multiple R-squared:  0.3479,    Adjusted R-squared:  0.2954
F-statistic: 6.631 on 7 and 87 DF,  p-value: 2.752e-06
```

- Higher r-sqrd than guards model
- All inputs significant

Model Performance: Big Men

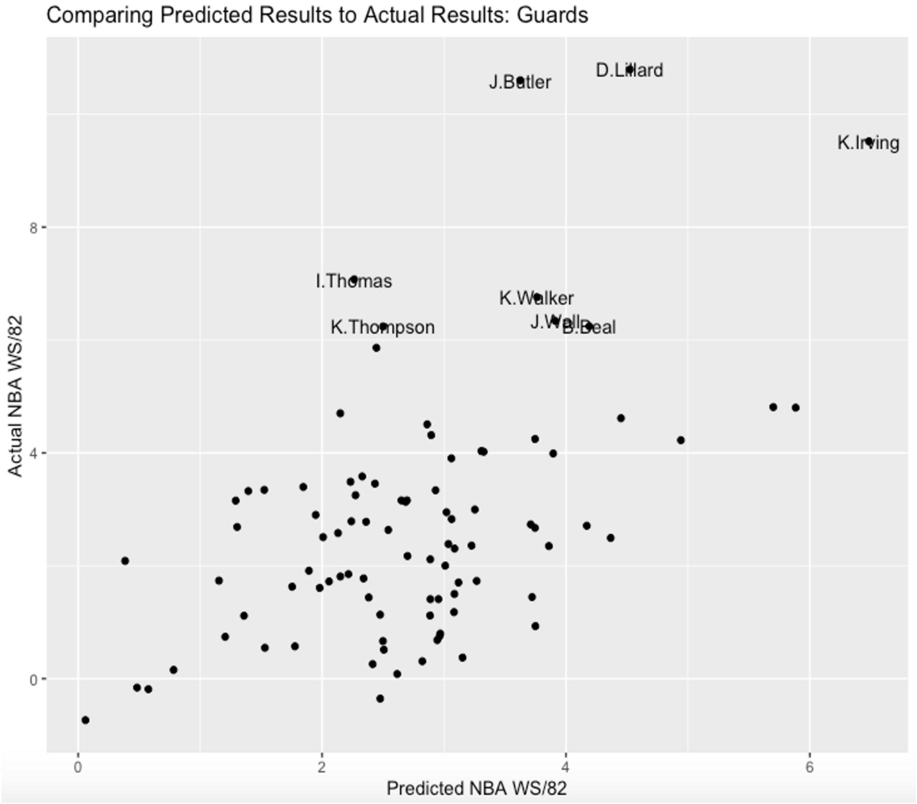


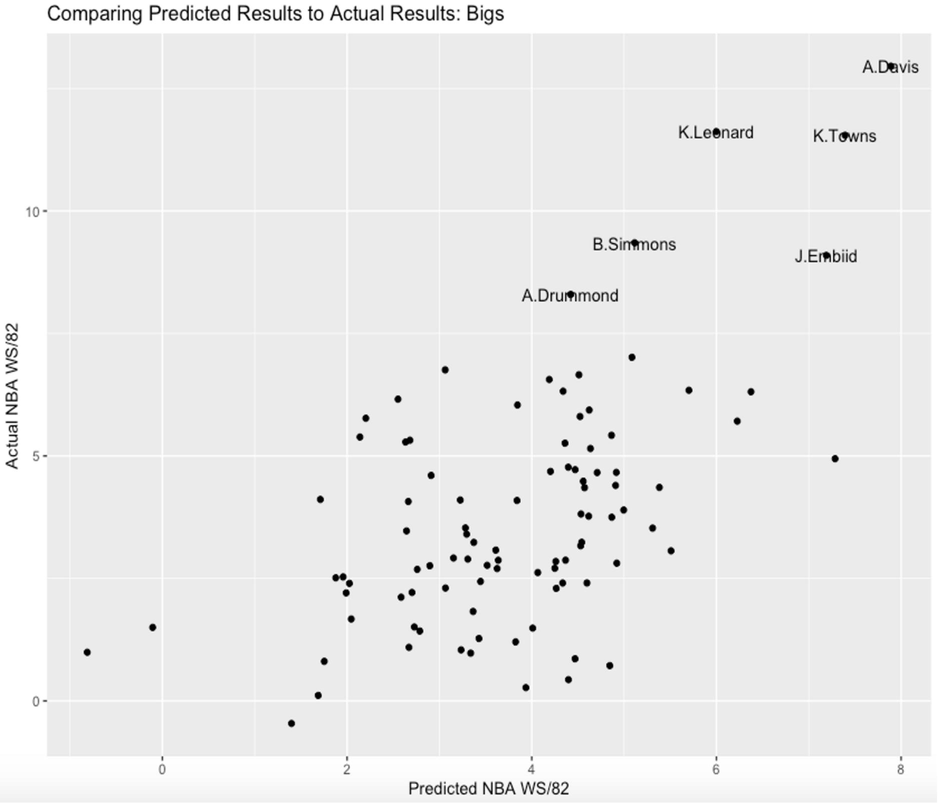
Top performing big men were all at or above the line

No players that were well below the line did exceptionally well

A Final Look

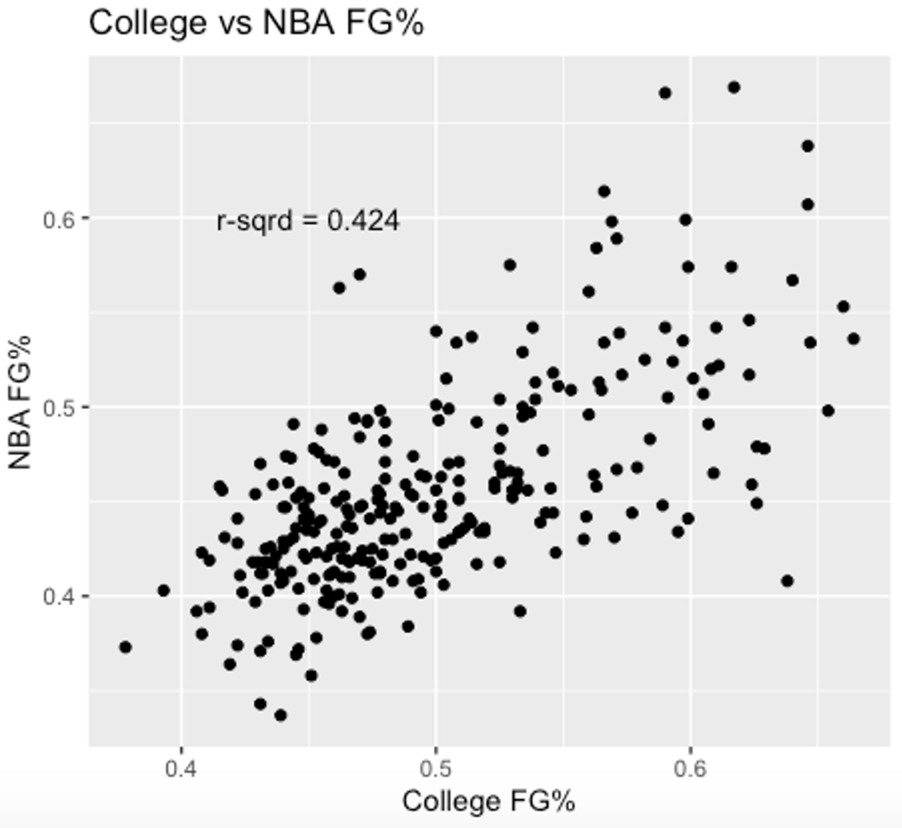
- There looks to be a moderate association for each group
- R-sqrd and plots show big men's careers easier to predict (though neither is very predictable)

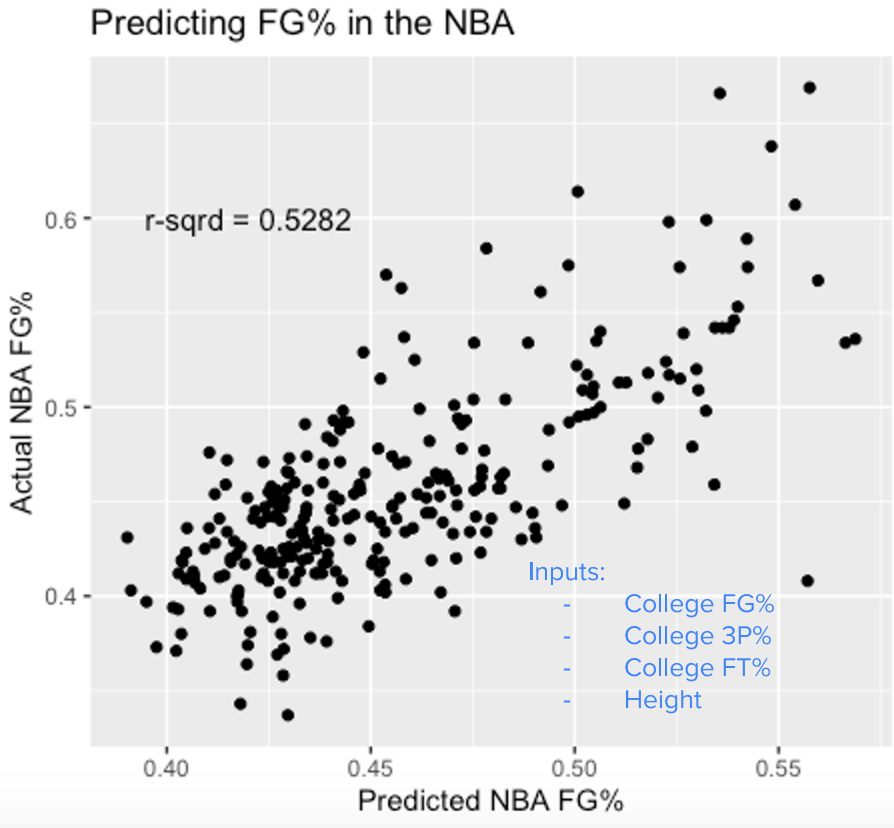




Predicting Future Shooting: FG%

Our group also tried to project future NBA shooting splits using players' college stats

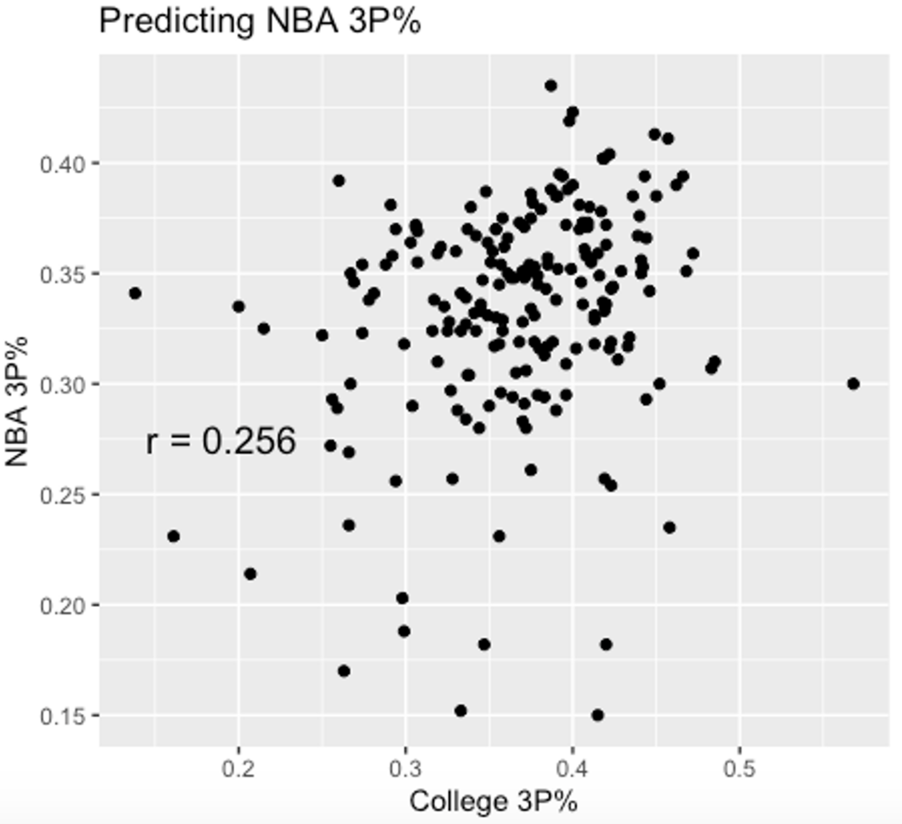


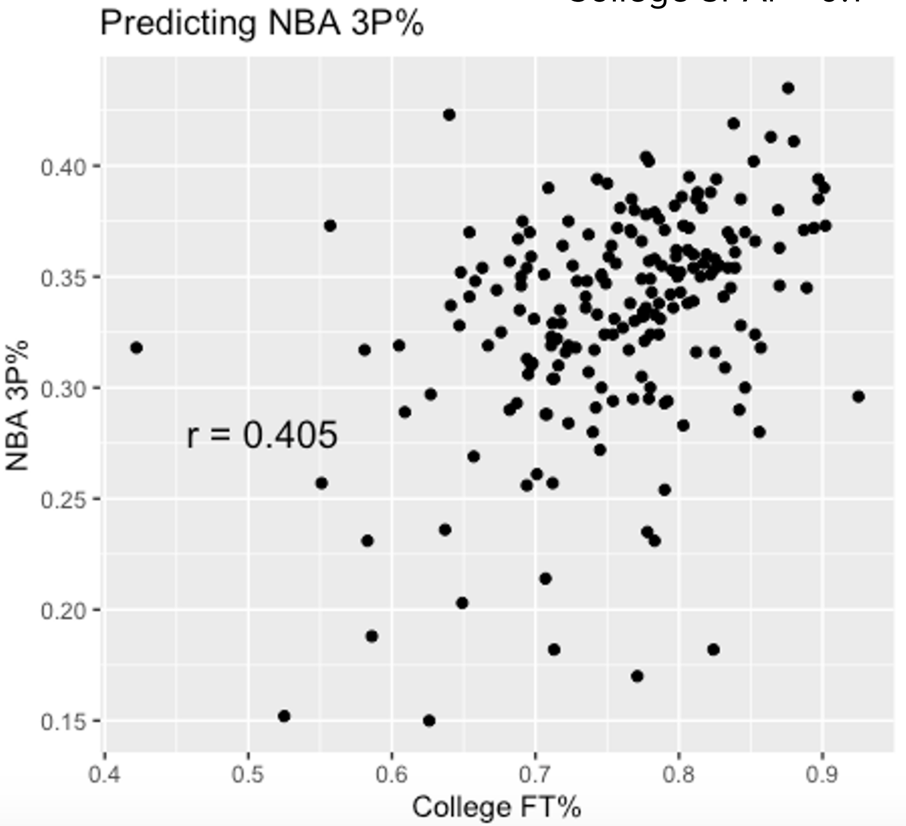


Predicting Future Shooting: 3P%

College FT% predicts NBA 3P% better than College 3P%

Note: Players had College 3PAr > 0.1





Predicting Future Shooting: 3P%

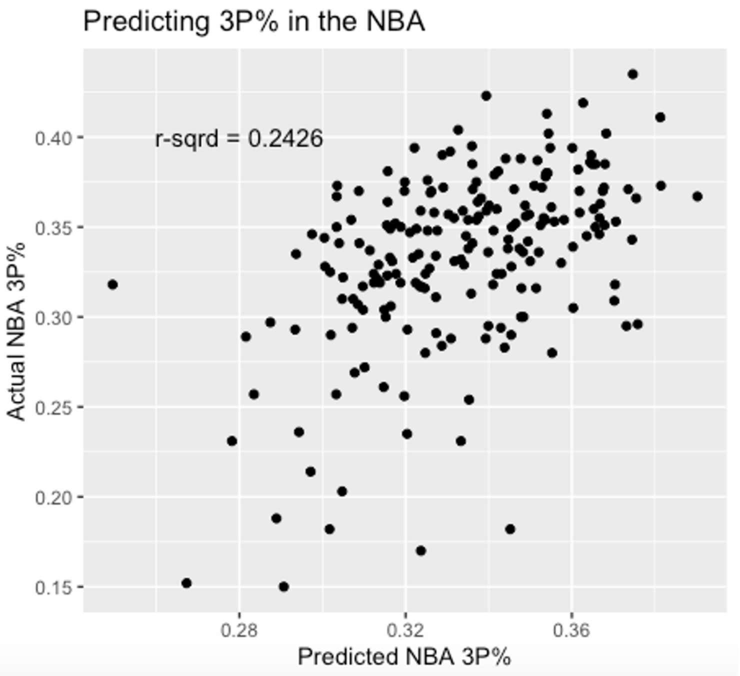
College 3P% was NOT a significant variable in a multiple linear regression for NBA 3P%

Inputs: College FT%, College 3PAr

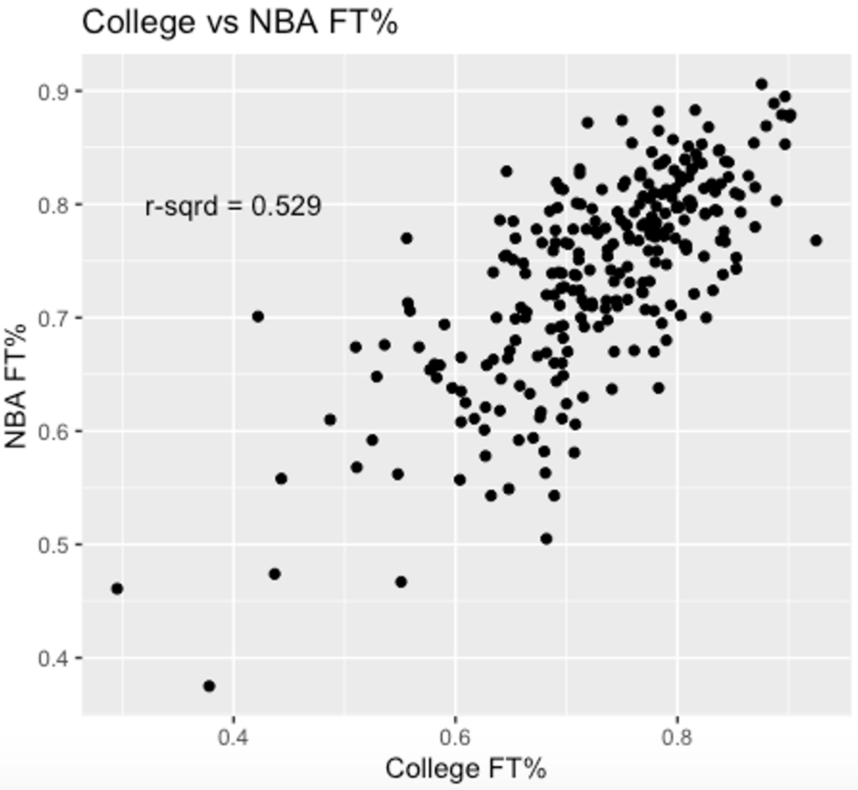
COMPARISONS

R-sqrd for univariate models:

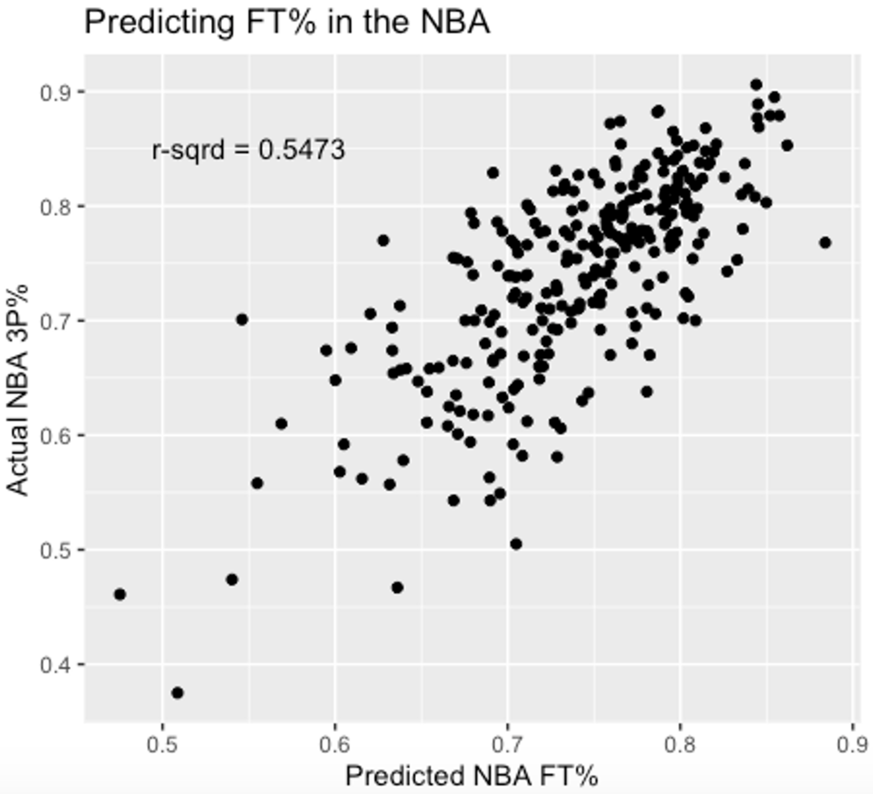
- College 3P%: 0.066
- College FT%: 0.164



Predicting Future Shooting: FT%



Inputs: College FT%, Height



Areas for Improvement

- While our models were adequate for predicting future NBA performance, they could have been improved
- Potential Issues / Methods for Improvement
    - Larger Sample Size
    - Potential Overfitting
    - Non-linear Fits
    - Non symmetric distribution of residuals
    - Bias towards Big Men

The Results of Our Model

- The best way to describe how our model did is to showcase some of the predictions that it got right (big hits) and some of the predictions that it got wrong (big misses).
- What makes a prediction a big hit or a big miss?
  - A big hit is a player that our model predicted to outperform or underperform his predicted win shares per season based on his draft position, and did so. A big miss is a player that our model predicted to outperform or underperform his predicted win shares per season based on draft position, and did not do so.

Which Statistics Did We Use to Determine Big Hits and Big Misses?

- A player would be able to be considered a hit or a miss if our model predicted them to either outperform or underperform their expected win shares based on draft position.
    - Therefore, we looked at the discrepancy between our model's predicted win shares per season for a player and the predicted win shares per season based on draft slot (predicted difference). If the difference between these two was large in either the negative or positive direction, then a player could be considered a hit or a miss.
    - As discussed earlier, if our model predicted a player to overperform or underperform their expected win shares per season based on draft position, and that player did so, that's a hit. If our model predicted a player to overperform or underperform their expected win shares per season based on draft position and the player did not do so, that's a miss.
    - For misses, we also looked at the discrepancy between a player's predicted win

Big Hits: Frontcourt

- Anthony Davis (predicted difference: 2.93 WS/82)
- Pascal Siakam (predicted difference: 2.81 WS/82, Difference between predicted win shares per season and actual win shares per season: -0.06 WS/82)
- Karl Anthony-Towns (predicted difference: 2.43 WS/82)
- Joel Embiid (predicted difference: 2.34 WS/82)
- Kawhi Leonard (predicted difference: 1.79 WS/82)
- Draymond Green (predicted difference: 1.20 WS/82)
- Anthony Bennett (predicted difference: -1.02 WS/82)

Big Hits: Backcourt

- Spencer Dinwiddie (predicted difference: 3.36 WS/82)
- Kyrie Irving (predicted difference: 2.60 WS/82)
- Delon Wright (predicted difference: 1.36 WS/82)
- Jimmy Butler (predicted difference: 0.95 WS/82)
- Damian Lillard (predicted difference: 0.85 WS/82)
- Isaiah Thomas (predicted difference: 0.84 WS/82)

Big Misses: Frontcourt

- Nerlens Noel (predicted difference: 2.59 WS/82, residual: -2.34 WS/82)
- Andre Drummond (predicted difference: -0.11 WS/82, residual: 3.87 WS/82)
- Otto Porter Jr. (predicted difference: -0.23 WS/82, residual: 1.31 WS/82)
- Harrison Barnes (predicted difference: -1.73 WS/82, residual: 1.70 WS/82)
- Jaylen Brown (predicted difference: -2.19 WS/82, residual: 1.41 WS/82)

Big Misses: Backcourt

- Buddy Hield (predicted difference: -1.44 WS/82, residual: 1.25 WS/82)
- Klay Thompson (predicted difference: -0.97 WS/82, residual: 3.74 WS/82)
- Jamal Murray (predicted difference: -0.77 WS/82, residual: 1.64 WS/82)

Applying Findings to the 2020 NBA Draft

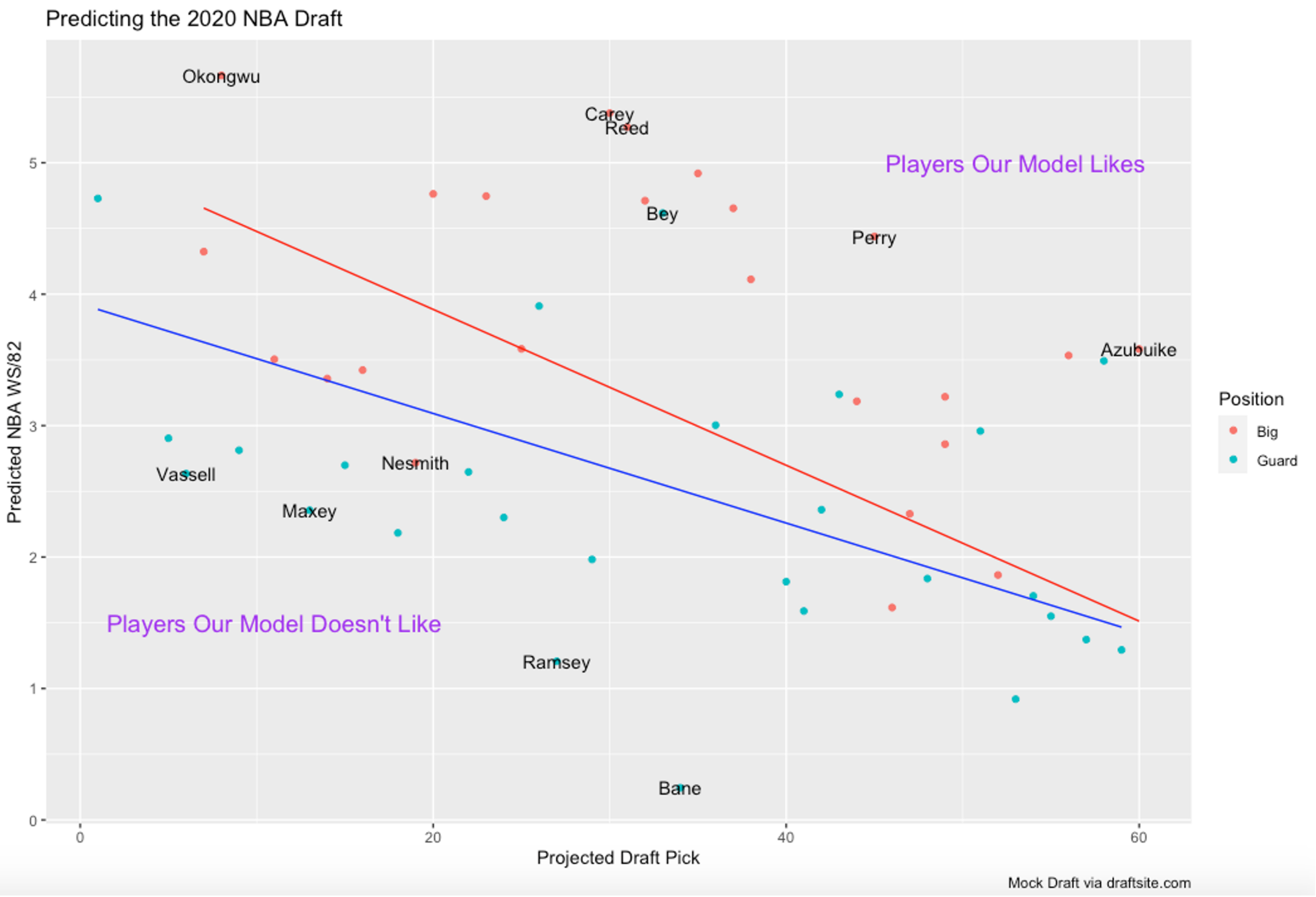
Who are the best players from this year's draft?

Mock Draft

We had to use a mock draft in order to compare results of our model versus the model using only draft picks

Website: draftsite.com/nba/mock-draft/2020/round1/

ROUND 1						
1	Golden State	Anthony Edwards	Shooting Guard	Georgia	6' 3"	225
2	Cleveland	James Wiseman	Center	Memphis	7' 1"	231
3	Minnesota	LaMelo Ball	Point Guard	United States	6' 6"	165
4	Atlanta	Killian Hayes	Point Guard	France	6' 5"	192
5	Detroit	Tyrese Haliburton	Point Guard	Iowa State	6' 5"	160
6	New York	Devin Vassell	Shooting Guard	Florida State	6' 5"	170
7	Chicago	Obidiah Toppin	Power Forward	Dayton	6' 9"	220
8	Charlotte	Onyeka Okongwu	Center	USC	6' 9"	234
9	Washington	Cole Anthony	Point Guard	UNC	6' 3"	184
10	Phoenix	Deni Avdija	Small Forward	Israel	6' 9"	215
11	San Antonio	Isaac Okoro	Small Forward	Auburn	6' 6"	213
12	Sacramento	RJ Hampton	Point Guard	New Zealand (NBL)	6' 5"	188



Predicting the 2020 NBA Draft
Okongwu
Carey
Reed
Bey
Perry
Players Our Model Likes
Azubuike
Nesmith
Vassell
Maxey
Players Our Model Doesn't Like
Ramsey
Bane
Predicted NBA WS/82
Projected Draft Pick
0
1
2
3
4
5
0
20
40
60
Position
Big
Guard
Mock Draft via draftsite.com

Top 10 College Players

Anthony Edwards:
5th highest Predicted NBA WS/82 for Guards in sample

	Player	Position	College	Mock Draft Pick	Predicted NBA WS/82
1	Onyeka Okongwu	Big	USC	8	5.664468
2	Vernon Carey	Big	Duke	30	5.377599
3	Paul Reed	Big	DePaul	31	5.271507
4	Isaiah Stewart	Big	Washington	35	4.919103
5	Jalen Smith	Big	Maryland	20	4.762780
6	Precious Achiuwa	Big	Memphis	23	4.746229
7	Anthony Edwards	Guard	Georgia	1	4.728330
8	Daniel Oturu	Big	Minnesota	32	4.711340
9	Xavier Tillman	Big	Michigan State	37	4.653649
10	Tyler Bey	Guard	Colorado	33	4.617134

Potential Bust and Sleeper Prospects

Player	Position	College	Mock Draft Pick	Predicted NBA WS/82	Difference
Vernon Carey	Big	Duke	30	5.377599	2.0863367
Udoka Azubuike	Big	Kansas	60	3.582787	2.0703968
Tyler Bey	Guard	Colorado	33	4.617134	2.0665517
Paul Reed	Big	DePaul	31	5.271507	2.0395402
Reggie Perry	Big	Mississippi State	45	4.439217	2.0373904
Myles Powell	Guard	Seton Hall	58	3.492682	1.9841888
Isaiah Stewart	Big	Washington	35	4.919103	1.9243195
Lamine Diane	Big	Cal State Northridge	56	3.533683	1.7841099
Xavier Tillman	Big	Michigan State	37	4.653649	1.7774568
Daniel Oturu	Big	Minnesota	32	4.711340	1.5386688

Difference: difference between projected NBA WS/82 based on model and projected NBA WS/82 based on draft pick and position

- \+ Difference = Underrated
- \- Difference = Overrated

Player	Position	College	Mock Draft Pick	Predicted NBA WS/82	Difference
Desmond Bane	Guard	TCU	34	0.244472	-2.2644264
Jahmius Ramsey	Guard	Texas Tech	27	1.207085	-1.5935981
Aaron Nesmith	Big	Vanderbilt	19	2.717723	-1.2257925
Devin Vassell	Guard	Florida State	6	2.634777	-1.0412608
Tyrese Maxey	Guard	Kentucky	13	2.356454	-1.0277989
Josh Green	Guard	Arizona	18	2.183898	-0.9919373
Isaac Okoro	Big	Auburn	11	3.504751	-0.9131312
Patrick Williams	Big	Florida State	14	3.357956	-0.8820391
Tyrese Haliburton	Guard	Iowa State	5	2.904087	-0.8136349
Nate Hinton	Guard	Houston	53	0.919154	-0.7977566

Potential Steals:
Carey, Azubuike, Bey, Reed, Perry

Potential Busts:
Nesmith, Vassell, Maxey, Green, Okoro

Best Shooters - Field Goal %

Player	Position	College	Mock Pick	Pred FG%	Pred 3P%	Pred FT%
Udoka Azubuike	Big	Kansas	60	0.6097612	0.2379255	0.5465026
Nick Richards	Big	Kentucky	49	0.5382756	0.2917133	0.7419836
Onyeka Okongwu	Big	USC	8	0.5107557	0.2879404	0.7283443
Obi Toppin	Big	Dayton	7	0.5103569	0.3097406	0.7172085
Vernon Carey	Big	Duke	30	0.4952074	0.2852066	0.6943326
Zeke Nnaji	Big	Arizona	38	0.4922704	0.2998915	0.7469329
Isaiah Stewart	Big	Washington	35	0.4883100	0.3029419	0.7617517
Daniel Oturu	Big	Minnesota	32	0.4874659	0.2994070	0.7172229
Xavier Tillman	Big	Michigan State	37	0.4872152	0.2980252	0.6986344
Jalen Smith	Big	Maryland	20	0.4738344	0.3264725	0.7438251

Udoka Azubuike

Best Shooters - 3 Point %

Player	Position	College	Mock Pick	Pred FG%	Pred 3P%	Pred FT%
Isaiah Joe	Guard	Arkansas	43	0.3841535	0.4117106	0.8458313
Tyshon Alexander	Guard	Creighton	54	0.4051052	0.3779599	0.8303504
Aaron Nesmith	Big	Vanderbilt	19	0.4369996	0.3749264	0.8025397
Tyrell Terry	Guard	Stanford	22	0.3974918	0.3731296	0.8587653
Malachi Flynn	Guard	San Diego State	42	0.4027342	0.3714014	0.8377310
Tyrese Haliburton	Guard	Iowa State	5	0.4378181	0.3677388	0.8037626
Trevelin Queen	Guard	New Mexico	55	0.4306197	0.3658519	0.7957345
Myles Powell	Guard	Seton Hall	58	0.3988055	0.3654598	0.7962954
Elijah Hughes	Big	Syracuse	44	0.4167088	0.3640432	0.7951159
Payton Pritchard	Guard	Oregon	59	0.4164258	0.3614005	0.8123805

Isaiah Joe’s 41.2% Predicted 3P% is higher than anyone from the sample set

The End