Predicting Career Success of Drafted Quarterbacks in the National Football League

Tam Nguyen
College of Wooster
Wooster, OH
tnguyen20@wooster.edu

Benjamin Kumpf
College of Wooster
Wooster, OH
bkumpf18@wooster.edu

Abstract

Since the 1983 Draft class Quarterbacks have been one of the most sought after players in hope of finding a born leader with the skills to become a franchise quarterback in the NFL. Their high draft priority has made the prediction of QB success from the draft a substantial topic of interest in the sports modeling community. More often then not, the players drafted high in a round often are much better in terms of career success, however outliars like Tom Brady and Russell Wilson defy this trend. By looking at college and combine statistics for these models show significant error with the accuracy of the model having about 30% error using various parameters to define success such as total games played in NFL, Average Value (AV), or Draft Pick (?). Since these models are assuming an inside source within an NFL team and have no knowledge of the team or draft position. Since there are no prior models including this knowledge, we wanted to determine if adding this information to the model could improve our accuracy. In result, coaching and team statistics did not improve the model. Our the best prediction metric was QB rating which had an error 25% less than our baseline and overall having a 13.16% error overall when testing against Quarterbacks drafted since 2008 when using combine and college statistics only. While trying to predict other metrics of success our model could not improve upon the baseline prediction which was the mean of the predicted value. This suggests the best predictive power using information available pre-draft is optimal at predicting for NFL Passer Rating.

Introduction

The NFL draft is often regarded as a make or break situation for the 32 teams trying to improve their team with fresh young talent. As the general of the offence, the quarterback (QB) is often among the most highly anticipated picks to change the destiny of the team. For example, in the 1998 NFL draft there were two quarterbacks who were considered equal talent: Peyton Manning and Ryan Leaf. The Indianapolis Colts were split between which QB they would select with their first pick until just hours before the selection, deciding on Peyton Manning the morning of the draft. The San Diego Chargers were considered by around 50% of the NFL to have gotten the best quarterback in the draft even though they got the second quarterback. However, the careers of these two diverged in major ways as Peyton Manning won the first Super Bowl for the Indianapolis Colts and set an NFL record for most valuable player awards while Ryan Leaf was cut by the Chargers within four years of being drafted. The stark contrast in Passer Rating with Ryan Leaf having a passer rating of 39 and Peyton Manning having a rating of 71(10). This shows the "hit-or-miss" nature of trying to draft a franchise NFL quarterback in the first round of the NFL draft. Sometimes teams find diamonds in the rough as few would have predicted a sixth-round pick, Tom Brady, would become one of the greatest QB's of all time.

This poses a question to be asked; what about the successful QB's sets them apart from other successful collegiate quarterbacks? Players such as potential hall of fame selections like Tom Brady can slip into the 7th round of the draft while Heisman winning and first round Johnny Manzel was not a success. In previous models predicting QB success have debated the importance of combine performance since there are no QB specific drills to gauge position-specific skill. Although these statistics can reflect general athleticism, they fail short at predicting what quarterbacks have the highest potential in the draft. Recent busts such as Johnny Manziel have had problems in their personal lives translate to their performance on the field. By incorporating college statistics, combine performance, and other existential factors such Super Bowl wins of the Drafting head coach, his playoff and regular season record, along with the offensive coordinator's (OC) overall winning percentage this could potentially clarify the QB prediction problem.

Previous Research

There are not many outstanding models to predict a quarterback's future performance. In 2008, Malcolm Gladwell drew attention to the difficulty of whether or not a college quarterback will be success in the NFL. Since a QB is a significant portion of a teams salary in a fixed-cap system, the maximum utility out of a young quarterback with a smaller contract could allow for the future success of a professional team. Various metrics are used to define QB success. In Adonna (2011) they used a binomial outcome for games played in the first season and a numeric value of net points contributed by the quarterback. The result of their model concluded that the quantitative statistics do not produce the most accurate prediction of a successful quarterback and thus qualitative factors, such as a scout's evaluation, could improve the clarity of the "quarterback prediction problem". Berri and Simmons(2009) (1), who concluded that the draft position of a quarterback had a considerable impact on how much that quarterback played, but not on how well he performed in the NFL. Increased playing time does correlate to players drafted in the first round since teams have more confidence in a higher value player. In 2010, Massey and Thaler focused on the correlation between compensation and player's performance (5). However, this work again showed that there was an overpaying for first-round draft and contract value did not correlate to success. (4).

If a precedence has been set by the previous attempts to predict quarterback success, it is elusive. As much if this work suggests, there is no clue for NFL teams to identifying college quarterbacks who are likely to succeed in the National Football Player career. Two possible explanations are:

1. NFL teams may de-emphasizing attributes which maybe be more significant such as the offensive coordinator's cumulative statistics statistics while emphasizing the strength of defense played against in college or throwing mechanics.
2. Factors determining quarterback's success are inherently biased by time, such as games played.

In this model, we try to solve both issues adding statistics to describe coaches of importance to a quarterback such as the head coach and offensive coordinator and by adding metrics to account for QB's who obtain the majority of their passing yards on short yard completions by combining the product of completion percentage and total passing yards. As a result of conflicting results stating that combine and college statistics are inconclusive at predicting QB success, we sought to determine which dependent variable is best predicted using pre-draft college and combine statistics.

Statistical data

Draft position, NFL statics, and college statistics were taken from Pro-Football-Reference.com for all drafted quarterbacks from 1987. Player's combine statistics back to 1987 seasons was obtained from Total-football-stats.com. For career college statistics for quarterbacks who do not have data, we looked up their data from college's website and several outside sources.

We believe that the drafted team rank should affect quarterback's performance, since they will meet better team's player and lead to a better result in upcoming season. Pulling NFL teams' rank, we based on the picking order of drafted round since, according to nfl.com.

All of this was pulled and written to Excel files. In the process of cleaning the data, there were many drafted QB's prior to 1987 without draft statistics and others with unavailable college statistics. Due to this reason those players were omitted from our data.

The last step before using the data as input to model was to average the combine statistics such as Broad Jump to fill in player's blank statistics. This final number was the one that was used as input to our model. These calculations were all done with simple Excel formula.

Variables

Dependant variable

In this model, we evaluated the variance among previous models variance in dependent variables chosen for success. Due to the variance of success metrics in previous models we chose to vary the dependant parameter, using Career Average Value, NFL Passer Rating, and Games played in the NFL.

Career AV the sum of the weighted seasons AV. Weighted season AV is based on the Approximate Value (AV) , which is an attempt to put a single number on the seasonal value of a player at any position from any year (since 1950). The formula for this variable was introduced by Doug Drinen in 1950. It is essentially the sum of the Approximate Value of each player for each season, with the weight of the best season contributing the most with the worst season contributing the least (). Therefore, CarAV is a good number to evaluate player's performance throughout his career with a number greater than 100 translating to a renown NFL quarterback. However, this dependent variable is biased with respect to time, and players with a longer career generally have a higher CarAV.

***CarAV** = 100%AV player's best season + 95%AV player's 2nd-best season + 90% of his 3rd-best season + 85% of his 4th-best season + ...*

In predicting Career AV we chose a Random Forest Regression and a Tree Classification model. The breakdown of CarAV as a binomial is listed below.

- Classify by binomial value:
  CarAV > 50 is 1
  CarAV < 50 is 0
- Classify by numerical value:
  Group 3: 0 ≥ CarAV ≥ 50
  Group 2: 50 < CarAV < 100
  Group 1: CarAV ≤ 100

By looking at the pattern of CarAV, we recognize that 50 is an adequate minimum threshold of success, since the average CarAV is 31. to determine player's performance since it can classify players into two different group of performance. Also, there are only a few player get a CarAV higher than 100 and they are classified as elite group or group 1.

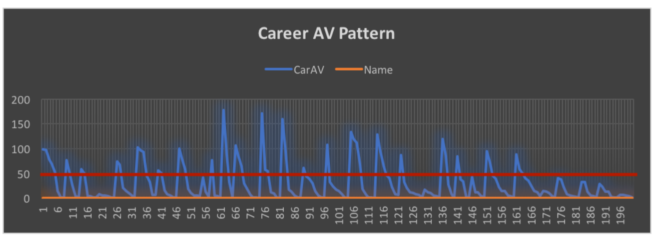

Figure 1: CarAV is plotted on the Y axis as a function of position in the data table

Quarterback Rating

A statistic attempting to determine the overall performance of a quarterback. The rating is made up of a series of calculations that take into account the number of completions, attempts, yards, touchdowns, and interceptions that a quarterback has recorded. The highest possible rating an NFL quarterback can have is 158.3; the lowest possible rating is 0. The formula is given below: (13)

- a = (((Comp/Att) * 100) - 30) / 20
- b = ((TDs/Att) * 100) / 5
- c = (9.5 - ((Int/Att) * 100)) / 4
- d = ((Yards/Att) - 3) / 4
- a, b, c and d can not be greater than 2.375 or less than zero.

QB Rating= (a + b + c + d) / .06
The maximum a passer can receive in any category from a to d is 2.375.

NFL Passer rating is a statistic that can be determined within the first year of a QB's career. This determination of success can be validated by the example of anticipated Hall of Fame quarterback Peyton Manning and Ryan Leaf who had a 71 and 39 as a Passer, respectfully, in their first professional year. Provided an accurate determination of Passer Rating this model could shed light on a QB's success in an NFL Offensive scheme. Also, this variable is only biased by time in the sense that a quarterback can either improve or worsen his Passer rating after his rookie year.

Games Played

Number of games that quarterback played through his career. The more games he played, the better chances to show his performance. Provided a successful model this could determine if pre-draft statistics and combine results predict player longevity.

Independent variables

In our analyses, we consider four cumulative statistic quantifying player's performance which are Combine statistics (player's index), College statistics, Drafted team statistics and Head Coach statistics.

1. **College statistics**
   G: Collegiate Games Played [Natural integers]
   Comp: Number of Completed Passes [Natural integers]
   Att: Number of Pass attempts [Natural integers]
   Yrds: Total Passing yards in College [Natural integers]
   Y/A: Passing Yards per passing attempt [Natural integers]
   AY/A: adjusted yards per passing attempt: (pass yards + 20*(pass TD) - 45*(interceptions thrown))/(passing attempts). ()
   TD: Passing Touchdowns in College [Natural intergers]
   Int: Total Interceptions Thrown [Natural intergers]
   Rate: QB Rating [Natural integers]
   C.Rank: Conference Rank of Collegiate Team
   Passing Efficiency: The equivalent metric to NFL Passer Rating but for college quarterbacks.
2. **Combine statistics**
   Ht: Players height [lbs]
   Wt: Players Weight [lbs]

Wonderlic: Score on Wonderlic Test
40yrd: Time in the 40 yard dash [s]
Vert: Vertical Jump Max Height [in]

BrdJ: Broad Jump Distance [in]

Shutt: Three cone shuttle time [s]

BMI: Body Mass Index

3. **Coaches statistics**
HCWP: Head Coach Winning percentage [%]
HCLP: Head Coach Losing percentages [%]
HCPO.W: Head Coach Play-offs winning percentages [%]
HCPO.L: Head Coach Play-offs loosing percentages [%]

4. **Drafted team statistics**
Round: Round Drafted [Natural integers]
Pick: Overall Pick in the Draft [Natural integers]
Year: Year drafted [Natural integers]
Drafted team rank: Drafted team rank at the current draft year [Natural integers]

Approach

Approaching the QB prediction problem, a model is required to predict the relationship of a given outcome using the factors mentioned prior. Therefore various modeling techniques and dependent variables were predicted to determine the optimal dependent variable for the factors obtained.

In order to have the rule to determine the future success of quarterbacks, making a decision tree was the primary method in predicting success. An advantage to using Tree Selection is its ability to assign specific values to problem, decisions, and outcomes of each decision. This reduces ambiguity in decision-making. Every possible scenario of a decision is considered and it defines a fork or node, enabling viewing all possible solutions to be viewed clearly. Therefore, these models could aid in the prediction of quarterback performance.

Since there are only have 200 players with available combine and college statistics, the data used was partitioned into a 70:30 training to testing ratio. The testing threshold was decreased below 80% in order to validate model against more established quarterbacks which were drafted after 2008. This gives us a maximum of 10 years experience to evaluate against. All models used 5-fold cross-validation. Since the binomial outcome is associated with the attribute of Career Average Value (CarAV), which is negatively biased towards players with a shorter career duration could skew model validation.

Our final approach focused solely on using the Random Forest regression to determine all of the sought after outcomes. In our model, we optimize a tuning parameter that governs the number of features that are randomly chosen to grow each tree from bootstrapped data. All our models are trained with 5-folds cross-validation to prevent overfitting. Model comparison was done by using the Mean Squared Error of each model and the average % error was calculated for NFL passer rating.

Results and Analysis

- **Model 1.1: Classifying future performance by binomial value**

  When trying to classify a player as a success or bust in terms of the binary CarAV predicted our error was significant. First, evaluating against players with less than 10 years experience is not a good method of evaluation and can lead to speculation of the models validity. When predicting CarAV several other factors are being predicted such as players career duration. This is problematic due to the inherent randomness of career ending injuries such as the most recent injury of Ryan Shazier this past season.

- **Model 1.2: Classify group of Carr AV**

  Since the binomial classification was speculative another approach was taken to classify the players into 3 groups. One drawback to this approach was the number of quarterbacks classified in the wrong group and the lack of specificity of these three thresholds. For example Robert Griffen III was had a CarAV of 45, however his career was cut short due to problematic knee issues which were reoccurring. By this metric Griffen would be considered a sub-par quarterback, however is career was cut short, therefore his CarAV could not be as high as another quarterback who had a longer career length.

- **Model 2: Random Forest Regression to predict different dependent variables**

  In a linear regression, the only factor deemed significant in the our dataset was overall pick number. Therefore, since the correlation in a linear model was poor, we chose take a Random Forest regression to predicting the numeric value of CarAV. This model did not perform better than the baseline as seen in figure 3 below. Therefore, we cannot predict the numeric CarAV beyond the mean of all CarAV's.

Prediction	Baseline MSE	MSE W/ Combine	MSE W/O Combine	MSE W/O Draft	Model Type
QB Rating	181.8	155.9	158.0	140.2	Random Forest
Games Played	1435.5	2413.1	2918.6	2970.7	Random Forest
Carr AV	633.0	1054.1	1068.8	1112.8	Random Forest
Draft Pick	4129.0	4333.0	4260.4	-------------	Random Forest

Figure 2: Error of our models using different factors and dependent variables

Another dependent we tested was the draft pick of an incoming quarterback. Again, based on the results of our model in Figure 3 our model was inconclusive and unable to improve beyond the baseline.

Third, the total games played in a players career was the dependent variable to be predicted. Upon the analysis of our model, we were again unsuccessful at predicting this outcome better than the baseline mean.

Finally, the best dependent variable predicted considering all factors was the NFL Passer Rating. Our lowest MSE and most improved model from the baseline included only college and combine statistics for the incoming quarterbacks. Having the lowest MSE and excluding information known post-draft such as round selected and overall pick, the best prediction of NFL Passer rating arises. A picture of the final tree is depicted in the figure 3 below.

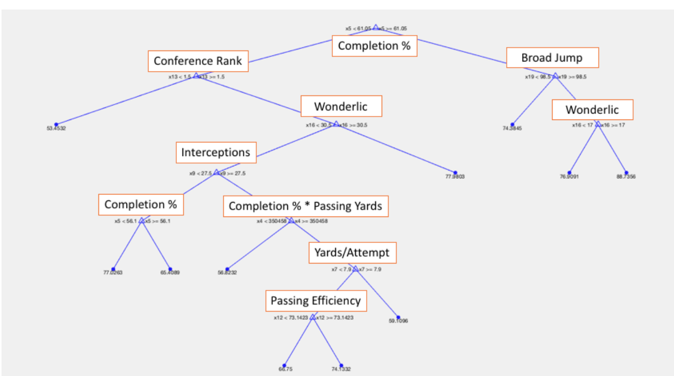

Figure 3: Tree produced by Random Forest to predict NFL passer Rating. The factors are labeled by white boxes played near the nodes.

In this tree the initial split is determined by a players college completion percentage. As seen in Figure 2, completion percentage, interceptions, Y/A, College Conference Rank, and Wonderlic are all significant in the tree with Wonderlic and completion percentage appearing multiple times in the tree.

The testing results for our model of NFL Passer Rating are shown below.

Player	Predicted	Actual Rating
Matthew Stafford	80.46	88.28
Sam Bradford	72.84	85.11
Colt McCoy	77.71	78.91
Tim Tebow	80.11	75.28
Jimmy Clausen	75.68	61.94
Tyrod Taylor	75.39	91.24
Colin Kaepernick	74.45	88.89
Andy Dalton	74.02	88.71
Cam Newton	68.96	85.34
Greg McElroy	68.94	79.23
Jake Locker	76.88	78.99
Christian Ponder	66.77	75.95
Blaine Gabbert	74.97	71.51
T.J. Yates	74.01	70.72
Ryan Mallett	64.22	66.76
Russell Wilson	77.10	98.76
Kirk Cousins	75.02	93.66
Robert Griffin	79.04	88.39
Nick Foles	76.19	87.44
Andrew Luck	72.80	87.33
Derek Carr	79.81	87.55
Blake Bortles	74.60	80.76
Johnny Manziel	70.00	74.35
Tom Savage	73.67	72.45
Dak Prescott	82.14	95.48

Figure 4: Results of Random Forest regression for NFL Passer Rating. The error associated with this model is 13.16%. Names in dark blue indicate good predictions where light blue indicates a poor prediction

Players such as Colt McCoy, Tim Tebow, Andrew Luck, and Cam Newton are not predicted well using this method. However, players which are good at passing from the pocket such as Matt Stafford, Kirk Cousins, Derek Carr, and Dak Prescott are predicted well. In terms of absolute evaluation, NFL Passer Rating is ambiguous. The average NFL Passer rating is 71, however a significant cluster of players with high CarAV tend to have a higher Passer rating. This can be visualized in the figure below. Therefore, a Passer rating of above a 71 should indicate success, while a value less than 71 should indicate a tendency for collegiate talent not to translate well into the NFL. This threshold is not absolute and has error associated with it since our model failed to predict the success of Cam Newton, a success in recent years. Despite the propensity for our model to predict lower Passer ratings than the true it is a bias of our model, which fails to predict the success of more dynamic and scrambling apt quarterbacks.

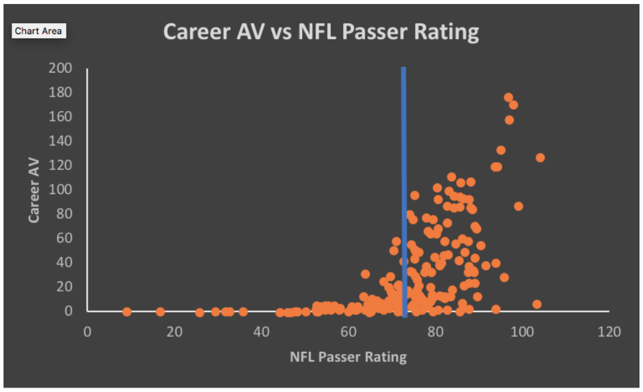

Figure 5: Career AV plotted against NFL Passer rating. The blue vertical line distinguishes the cluster of low CarAV and high CarAV

Although the predicted value of NFL Passer Rating is not perfect, the general indication of a higher passer rating could infer success. Since the main role of a quarterback is his passing ability, particularly across the line of scrimmage, therefore, a successful quarterback must have a strong affinity for passing. The players who occupy the top 2 % of CarAV have a passer rating which is 85 or greater.

2018 QB Draft Class	Predicted NFL Passer Rating
Baker Mayfield	80.17
Sam Darnold	78.56
Josh Allen	68.98
Josh Rosen	74.86
Lamar Jackson	65.62
Mason Rudolph	80.73
Kyle Lauletta	77.32
Mike White	79.27
Luke Falk	80.86
Tanner Lee	69.62
Danny Etling	70.65
Alex McGough	71.93
Logan Woodside	73.04

Figure 6: Predicting NFL Passer Rating of the 2018 Quarterback Draft Class

Using the Random Forrest approach we achieved the lowest observed error for predicting quarterback career success. Using the threshold of 0.45 to separate success or bust the model predicted the career fate of the following quarterbacks.

Our model predicts that Luke Falk, Mason Rudolph, and Baker Mayfield will be the best quarterbacks of that class. Lamar Jackson was deemed the biggest risk in the draft as he had the lowest predicted Passer rating.

Interpretation of the Model

The significance of NFL success metrics in previous quarterback predicting models have shown that quantitative statistics pre-draft have a tendency to miss significant first round busts. By using factors pre and post draft the only dependent variable that was predicted better than the baseline was NFL Passer rating. Like the NFL, college statisticians maintain a Passing Efficiency statistic which determines a quarterbacks production beyond total touchdowns and total passing yards. Having a similar statistic in the NFL allows for a greater correlation between the two, hence why we can predict NFL Passer rating much better than CarAV, total games, or draft pick. Although this is not the best factor to define success it is the best outcome by which college and combine data can predict.

In building several models to test many different outcomes we can conclude with Adonna (2011) in that NFL teams use qualitative data in conjunction with quantitative data efficiently. From the lack of predictive ability of draft round for a model trained on college and combine statistics indicates other factors not included in this model are more significant in successful quarterbacks.

Furthermore, the bias of our model to favor Quarterbacks who have an affinity for pocket passing indicates the changing dynamic of a successful quarterback through time is challenging to predict. Players with the athletic ability such as Cam Newton can extend plays with his feet and find an open passer that a traditional quarterback could not. Although combine statistics gauge general athletic ability they are not necessarily the best attributes to determine the quarterback specific athleticism.

The addition of the factor completion percentage times the total yards appeared in several trees during the process of modeling NFL Passer Rating. This rule in the tree indicates that players who have a high completion percentage and high total passing yards can be distinguished based on the Y/A and Passing efficiency. We see this as the tree's attempt to classify players who rely on short completions such as screen passes to be successful. Despite teams having different offensive styles, generally quarterbacks who have the best throwing ability down the field tend to be more successful in the NFL.

Lastly, to improve upon our model we propose a season by season prediction of NFL Passer rating would be the optimal way to model this outcome. By including game by game statistics the QB's performance in college against defenses who are good at stopping the passing game can be accounted for. Since the talent in the NFL far exceeds the collegiate level, a quarterbacks performance against an elite opponent should weigh considerably in the future success of a college quarterback.

Summary and Future Work

In conclusion, the only model we could improve upon the baseline was the prediction of NFL Passer Rating. Since this has a positive correlation with CarAV, it is a good indication of future success for a quarterback who is apt in the pocket. Failing to improve upon the baseline model for predicting CarAV, total games, and draft pick indicates quantitative statistics from a quarterbacks past are best at determining his Passer rating in the NFL rather than a career success metric. Factors such as injury, the presence of an established starting quarterback, and fit within a teams offensive scheme are all outside of the scope of this model. However, by improving the NFL Passer rating model we can see that college and combine statistics do correlate to an NFL metric, NFL Passer rating, with factors such as Wonderlic, Conference Strength, Combine scores, and Passing attributes playing an important role in the Random Forest regression. The model we did allow us to see which factors are the most predictive of an NFL metric and the limitations to using NFL Passer Rating and other dependant variables to classify success. Even when the future performance of drafted quarterbacks is inherently difficult, our model indicated that combine and college statistics are reasonable independents to predict the future NFL Passer rating to a 13.16% error. This model is only a small step in clarifying the quarterback prediction problem, but a significant one none the less.

References

[1] Berri, David J., and Rob Simmons. "Catching a Draft: On the Process of Selecting Quarterbacks in the National Football League Amateur Draft." Journal of Productivity Analysis 35, no. 1 (February 1, 2011): 37–49. https://doi.org/10.1007/s11123-009-0154-6.

[2] Julian, Wolfson, Addona Vittoria, and Schmicker Robert. "The Quarterback Prediction Problem: Forecasting the Performance of College Quarterbacks Selected in the NFL Draft." ResearchGate. Accessed May 8, 2018. https://www.researchgate.net/publication/220033161_The_Quarterback_Prediction_Problem_Forecasting_the_Performance_of_College_Quarterbacks_Selected_in_the_NFL_Draft.

[3] Kinsley, John R. "Jacoby Brissett Is One of the Most Underrated Quarterbacks in Football." Music City Miracles, April 4, 2018. https://www.musiccitymiracles.com/2018/4/4/17195570

[4] "Highly Paid NFL Player Kirk Cousins Still Drives a Dented Van." Accessed April 24, 2018. https://www.cnbc.com/2018/03/13/kirk-cousins-still-drives-a-dented-van.html

[5] Massey, Cade, and Richard Thaler. "Overconfidence vs. Market Efficiency in the National Football League." Working Paper. National Bureau of Economic Research, April 2005. https://doi.org/10.3386/w11270.

[6] J.D. Pitts & B. Evans (2017) Evidence on the importance of cognitive ability tests for NFL quarterbacks: what are the relationships among Wonderlic scores, draft positions and NFL performance outcomes?, Applied Economics, 50:27, 2957-2966, DOI: 10.1080/00036846.2017.1412081

[7] Kuzmits, F. E., & Adams, A. J. (2008). THE NFL COMBINE: DOES IT PREDICT PERFORMANCE IN THE NATIONAL FOOTBALL LEAGUE? Journal of Strength and Conditioning Research, 22(6), 1721-7. Retrieved from https://search.proquest.com/docview/213057745?accountid=15131

[8] McKenzie, G. (2015). Data analytics in sports: Improving the accuracy of NFL draft selection using supervised learning (Order No. 10182549). Available from ProQuest Dissertations & Theses A&I. (1842666793). Retrieved from https://search.proquest.com/docview/1842666793?accountid=15131

[9] Mahoney, J. (2017, June 28). Success rates of drafted quarterbacks - Mile High Report. Retrieved March 28, 2018, from https://www.milehighreport.com/2017/6/28/15880748/success-rates-of-drafted-quarterbacks

[10] Mulholland, J., & Jensen, S. T. (2014). Predicting the draft and career success of tight ends in the National Football League. Journal of Quantitative Analysis in Sports, 10(4). https://doi.org/10.1515/jqas-2013-0134

[11] NFL and AFL Draft History. (n.d.). Retrieved March 28, 2018, from https://www.pro-football-reference.com/draft/

[12] Wallace Hendricks, Lawrence DeBrock & Roger Koenker, "Uncertainty, Hiring, and Subsequent Performance: The NFL Draft," Journal of Labor Economics 21, no. 4 (October 2003): 857-886.

[13] "Quarterback Rating - QBRat Definition - Sporting Charts." Accessed May 8, 2018. /dictionary/nfl/quarterback-rating-qbrat/.