Drafting NFL Wide Receivers: Hit or Miss?

By Amrit Dhar

## I. Introduction

The Detroit Lions, an NFL franchise known for regularly fielding poor football teams, attained a cumulative win/loss record of 48-128 from the 2000-2010 seasons. Many football analysts believe that part of their failure to create quality football teams is due to their aggression in selecting wide receivers early in the NFL draft, and their inability to accurately choose wide receivers that become elite NFL players. Over the past decade, they have spent four of their 1st round draft picks on wide receivers, and only two of those picks actually remained with the Lions for more than two years. The Lions represent an extreme example, but do highlight the inherent unpredictability in drafting wide receivers that perform well in the NFL. However, teams continue to draft wide receivers in the 1st round like the Lions have done as the NFL has evolved into a "passing" league. In 2010 alone, 59 percent of NFL play-calls were called passes, which explains the need for elite wide receivers in any franchise.

In this report, I want to analyze whether the factors that teams believe are indicative of wide receiver effectiveness in the NFL actually do lead to higher performance. The above anecdote suggests that there is a gap between how NFL teams value wide receivers in the draft and how well they perform in the NFL. By conducting statistical analyses of where wide receivers were chosen in the NFL draft against how they performed in the NFL, I will be able to determine some important factors that have lead to their success in the NFL, and will be able to see whether those factors correspond to the factors that NFL draft evaluators believe are important for success in the NFL. The rest of this report will continue as follows: **II. Data Description** will discuss the source and type of data used in the analysis, **III. Methods** will specify the statistical techniques used in both analyses, **IV. Draft Pick Prediction** will describe the regression analyses used for predicting the NFL draft position of a wide receiver, **V. NFL Performance Prediction** will detail the regression analyses used for predicting NFL wide receiver performance, and **VI. Conclusions** will discuss the most important ideas and insights to take from this report.

## II. Data Description

In doing these statistical analyses, I want to be able to figure out whether the information available to an NFL talent evaluator at the time of the draft can accurately predict NFL performance. Thus, I will only focus on data that is measured prior to the draft or is known to the evaluator before the player plays in the NFL. Talent evaluators generally look at two different sources of data when assessing players: their NFL combine results and their collegiate statistics. The NFL combine is a week-long event, occurring every February, where college football players perform physical and mental tests in front of NFL coaches, general managers, and talent evaluators. A player's performance at the combine can greatly shape the perception of their value going into the draft. Collegiate statistics for receivers generally consist of the number of receptions, receiving yards, and touchdowns over their entire career. Both sets of data will hopefully be useful in explaining wide receiver performance in the NFL and draft position.

Having considered the type of data I needed for my analysis, I then scraped the Draft Round, Pick Number, Team, Name, NFL receiving statistics, and College Name for all the 1999-2010 wide receiver draft picks, from pro-football-reference.com. There were a total of 377 observations in this span. I decided to use the 1999-2008 data as my training set for building the models, and chose to leave the 2009-2010 data as a test set for future research. I also included a dummy variable in my data matrix to indicate whether or not a prospect attended a BCS-conference school. This distinction generally is made for schools that have large student populations and have more scholarship money to give to collegiate athletes. These schools are also generally characterized as playing more competitive, quality football as compared to non-BCS schools. This leads to a large amount of media coverage given to BCS schools, which may impact the draft positions of wide receivers from non-BCS schools.

I also scraped the college receiving statistics for these 377 observations from totalfootballstats.com and www.sports-reference.com/cfb/. The data variables included were: their total college receptions, receiving yards, and touchdowns, as well as their final collegiate year's receptions, receiving yards, and touchdowns. The reason I scraped information on a player's final college season is because I hypothesized that players that do extraordinarily well in their final season while having generally average college careers, get overrated in the draft rankings and end up being poor receivers in the

NFL. I used these receiving statistics to create 2 new variables: the percentage of total receiving yards amassed in the final college season, and the percentage of total touchdowns amassed in the final college season.

To acquire NFL combine statistics, I scraped data from nflcombineresults.com, a website that has stored combine information for all drafted players from 1999-2011. Specifically, I obtained data for a player's Height/Weight (used to calculate BMI), his 40-yard dash time, his vertical leap, his broad jump (i.e. standing long jump), his 20 yard shuttle time, and his 3 cone drill time. BMI is otherwise known as Body Mass Index and is used as a metric to gauge a player's muscle per inch$^2$. In general society, high BMI's are indicative of obesity, but in the NFL most players have little body fat so BMI is used as a measurement of lean muscle mass. The shuttle and 3 cone drills are primarily used to measure one's agility and coordination while the first 3 combine events are used to measure pure speed, leaping ability, and explosiveness respectively. There are other combine events that players attend, but due to the lack of data available on the internet, I decided to only scrape data for these 5 events.

I also web-scraped the team Quarterback Rating (QB Rating) for all 32 NFL teams over the span of 1998-2010 from NFL.com. QB Rating is a measure of quarterback effectiveness, and I assigned the previous year's team QB rating to a wide receiver drafted by that team in the present. For example, if Player A was drafted by the Detroit Lions in 2011 and the Lions had a team QB rating of 88.5 in 2010, then Player A would be assigned 88.5 for that variable. If a player was drafted by a brand new NFL team, I imputed the current draft year's team QB rating instead (as there is no previous history). I included this information in the dataset, because it is often argued that great quarterbacks help make great wide receivers, and the most up-to-date, concise piece of information a talent evaluator has on his own team's quarterback situation leading up to the draft is the previous year's QB rating.

After accumulating all the data for the 377 drafted wide receivers in my sample, I decided to remove those observations that didn't have a height/weight measurement from the combine as all those observations didn't have any combine drill data as well and wouldn't have worked as suitable observations for my analysis. I also chose to focus my analysis on pure wide receivers so I removed NFL wide receivers that were either quarterbacks or running backs in college; this data removal process left me with 333 observations. 109 of the 333 observations had at least one missing value for the combine data, so I imputed those values with the 10 "nearest neighbors" median value for that specific variable, where the "nearest neighbors" were defined by the 10 observations that had the smallest Euclidean distance in normalized height/weight from the original missing data observation. I applied this procedure because I expected athletes with similar height and weight measurements to perform similarly in the combine events.

## III. Methods

For my two analyses, I estimate both parametric and non-parametric models in order to create a two-way comparison of my results and provide further evidence for any important predictors of draft position and/or NFL performance that I may find.

The parametric model used in my analyses is the standard ordinary least squares (OLS) model. For "*p*" predictor variables and a response variable *Y*, the model assumes that *Y* is a linear function of the model parameters $\beta_0$, $\beta_1$, $\beta_2$,…., $\beta_p$. To allow for statistical inference of the estimated coefficients, the error term is also assumed to be normally distributed. The primary benefits of ordinary least squares stem from its analytic simplicity. It is easy to interpret the estimated coefficients of an OLS model, by saying that an increase in predictor $X_p$ by one unit increases the estimated $E(Y \mid X_1, X_2, \ldots)$ by $\hat{\beta}_P$. One can often take transformations (i.e. natural log) of the predictor variables while still maintaining the parameter linearity assumption. In addition because of the small sample sizes in my dataset, OLS models may perform better because of its simple structure. However, the assumption of error normality may not be entirely appropriate for the dataset. Although OLS can model single variable effects well, it is not as suitable for modeling joint variable interactions.

The non-parametric model I decided to implement in my analyses was a recursive partitioning regression tree (also known as CART). CART models divide the "*p*"-dimensional predictor space into "*p*"-dimensional hyper-rectangles, and fit a constant in each rectangular region. The fitted constant is the average of the response variable for those observations that are in a particular rectangular region. A "greedy" algorithm chooses the binary variable splits for each rectangle sequentially by choosing a splitting variable (i.e. $X_j$) and split point (i.e. $X_j > m$ & $X_j < m$) that minimizes the sum of the squared errors in both of the split rectangular regions. The tree partitioning starts with all the observations at the root node (i.e. top of the tree) and stops when the number of observations in each rectangular region or leaf node is below a user-defined threshold. A huge advantage to implementing regression trees is the ease of interpretation; the visual tree structure is very simple to follow and it allows for joint variable interactions not easily constructed by OLS. In the context of my analyses, trees may be useful in grouping wide receivers that have similar characteristics together so one can see what wide receiver segments perform relatively better or worse in the NFL and rank high or low in the draft. More details about CART models can be found in (Elements of Statistical Learning, Freedman et al.).

## IV. Draft Pick Prediction

For my draft pick regression analysis, I excluded all the remaining observations that had no college data available online, as the collegiate statistics form a big chunk of my predictor variables for this analysis. This exclusion left me with 266 observations. **Figure 1** and **Figure 2** display the empirical distributions of the draft picks by draft round and pick number respectively for the 266 observations. The NFL draft has seven rounds, and in each round there are, on average, 35 picks per round.

Figure 1: Distribution of Draft Picks by Round

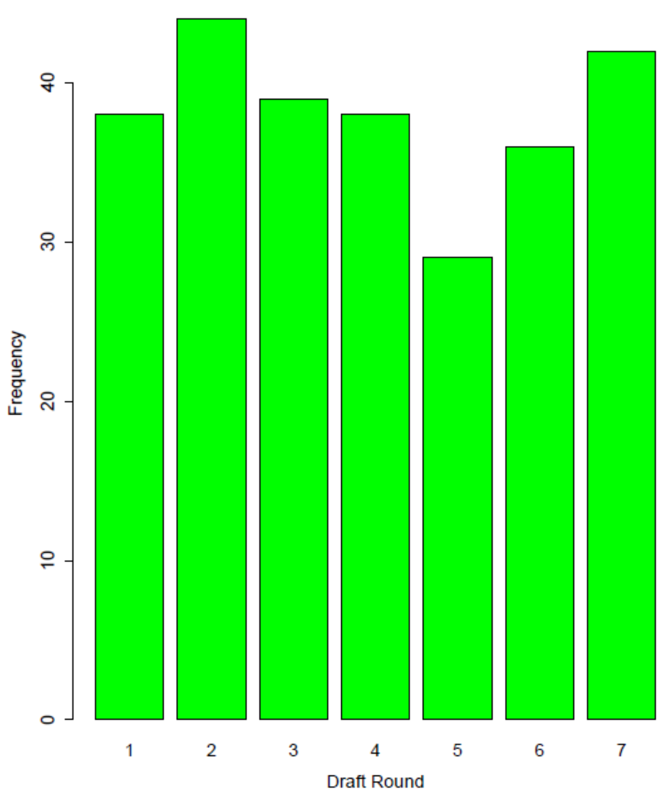
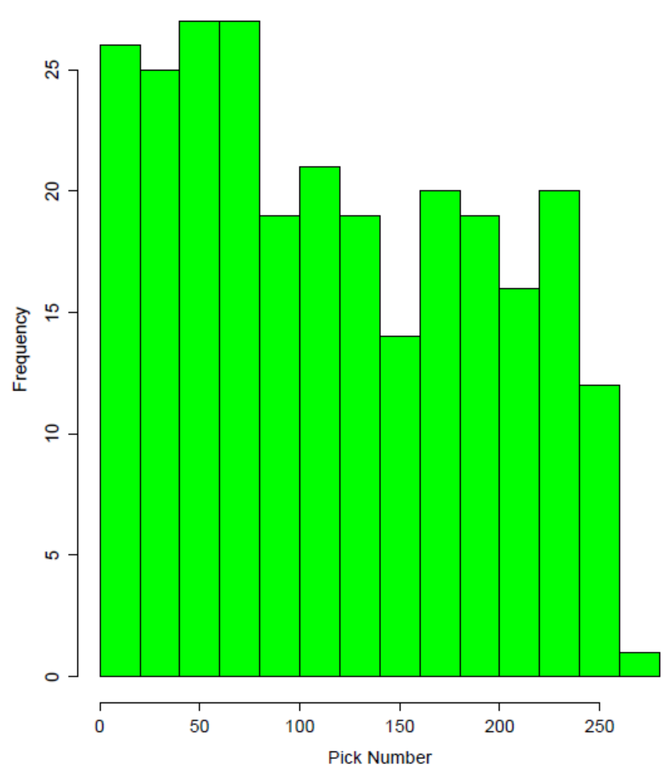


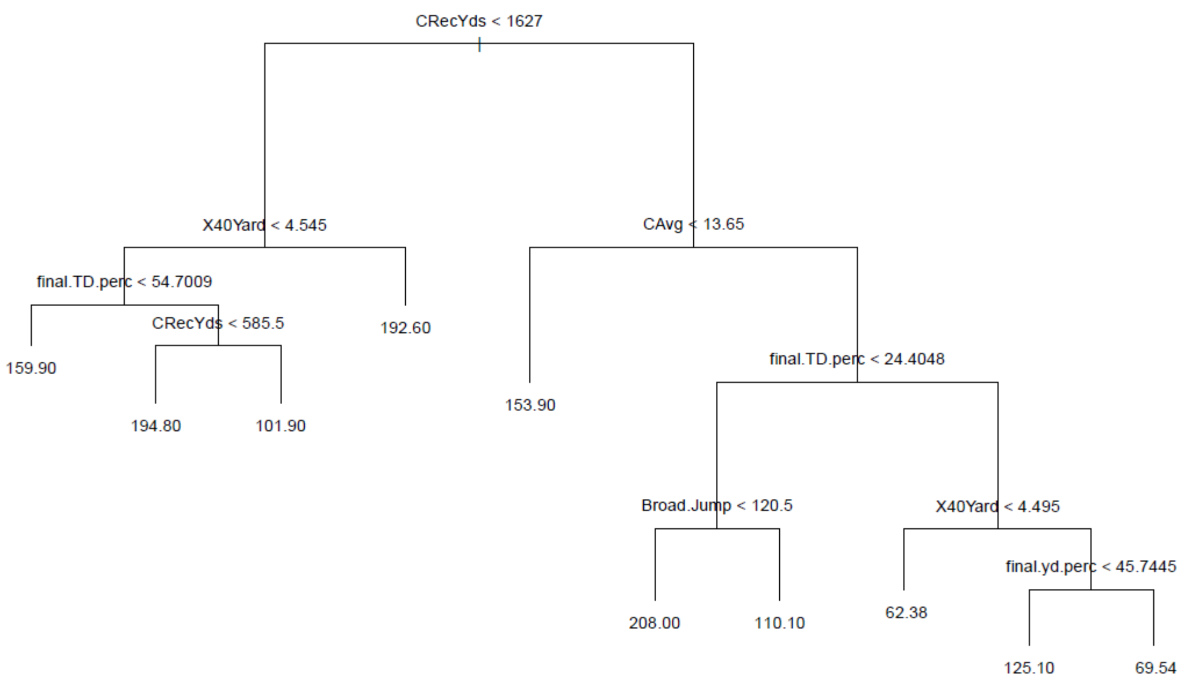

Figure 2: Distribution of Draft Picks by Pick Number



Looking at both **Figure 1** and **Figure 2**, there seems to be an adequate amount of variation throughout all the rounds and pick numbers. Thus, it seems appropriate to proceed with the OLS estimation and CART model fitting.

I first discuss the results from my tree regression analysis, and then introduce my results from the standard linear regression model. **Figure 3** shows the regression tree that was generated for predicting the draft ranking of a particular wide receiver from 1999-2008. The first thing to notice is that the Total Collegiate Receiving Yards (*CRecYds*) variable was chosen to be the first variable split. Mathematically, that variable was chosen because it was the variable that minimized the sum of squared errors at the two resulting nodes. Intuitively, it suggests that receivers with college receiving yards above or below the 1627 yard threshold are grouped into different categories, namely "Over-achievers" and "Under-achievers." I created these group names because any college receiver with less than 1627 receiving yards either played only a couple years of college football or just did not perform well in college. On the other hand, I believe receivers with college receiving yards above the denoted threshold are more likely to be experienced college football players and have had longer college careers. Looking at the terminal nodes on either side of the first variable split, it appears that these "Over-achievers", on average, are predicted to be drafted high than the "Under-achievers" (Note: high draft positions correspond to the smaller draft pick numbers).

**Figure 3: Pruned Regression Tree – Draft Pick Prediction**



From within the "Under-achievers" segment, the tree suggests that a wide receiver can rebound in the draft rankings (i.e. obtain a prediction near the approximate $102^{nd}$ draft pick) by running a fast 40-yard dash, by obtaining a little more than half of his total collegiate touchdowns in his final year, and by not having an abysmal amount of college receiving yards (i.e. total receiving yards less than 585). This 40-yard dash result is consistent with the notion that NFL talent evaluators tend to rank fast wide receivers very high, regardless of whether they have had strong college careers or not. However, the average 40-yard dash time is the sample was 4.48 so the "Under-achievers" seemed to not have had to run extremely fast times at the combine in order to get drafted in respectable positions. The observation that a high proportion of total college touchdowns in one's final season tends to lead to better draft predictions also makes sense because if a receiver has under-achieved for the majority of his college career, he will need an extraordinary final season to garner attention from NFL scouts. Another somewhat obvious insight is that slow receivers, who don't have stellar final seasons, tend to be drafted very low (as can be seen by the three other terminal nodes in the "Under-achievers" segment).

The "Over-achievers" portion of the tree suggests that high draft picks are characterized by players that have accumulated a large number of receiving yards, had a high yard per reception figure, ran an above-average 40-yard dash time, and had a moderate final season touchdown percentage. The results for this segment are fairly similar to the results found previously for "Under-achievers." One thing to notice is that the cutoff for the final year touchdown percentage (*final.TD.perc*) is at 24.4%, much less than the cutoff constructed for the "Under-achievers" at 54.7%. This presents evidence for the notion that experienced collegiate wide receivers tend not to have to have amazing final seasons in order to be drafted high.

My linear regression model uncovers some new insights about wide receiver draft pick rankings that were not explicitly expressed in the tree regression. NFL observers often pose the question of whether collegiate statistics or combine results are more indicative of NFL performance. I will use $R^2$ values from the linear regression models to gauge the explanatory power of both college and combine data on draft rankings and then mimic that in a similar analysis of NFL performance. Regressing the draft pick (response variable) only on the college related statistics gives me an $R^2$ of

0.2222 while the regression only on combine related variables receives an $R^2$ of 0.05445. OLS seems to suggest that collegiate statistics explain a higher proportion of the variation in draft rankings than do NFL combine drills (by a factor of approximately 4). My final OLS model for predicting draft position is shown in **Figure 4**.

**Figure 4: Draft Pick Prediction – Final OLS Model**

Coefficients:

| | Estimate | Std. Error | t value | Pr(>\|t\|) | |
|---|---|---|---|---|---|
| (Intercept) | -5.139e+02 | 2.010e+02 | -2.557 | 0.01114 | * |
| BCS | -4.020e+01 | 9.048e+00 | -4.443 | 1.32e-05 | *** |
| CRecYds | -5.514e-02 | 1.944e-02 | -2.836 | 0.00493 | ** |
| I(CRecYds^2) | 9.005e-06 | 4.185e-06 | 2.152 | 0.03235 | * |
| CAvg | -4.010e+00 | 1.701e+00 | -2.357 | 0.01917 | * |
| CRecTD | -2.372e+00 | 7.323e-01 | -3.239 | 0.00136 | ** |
| final.yd.perc | -9.249e-01 | 2.623e-01 | -3.526 | 0.00050 | *** |
| BMI | -5.655e+00 | 2.756e+00 | -2.052 | 0.04119 | * |
| X40Yard | 2.275e+02 | 4.266e+01 | 5.332 | 2.13e-07 | *** |

Adjusted R-squared: 0.302

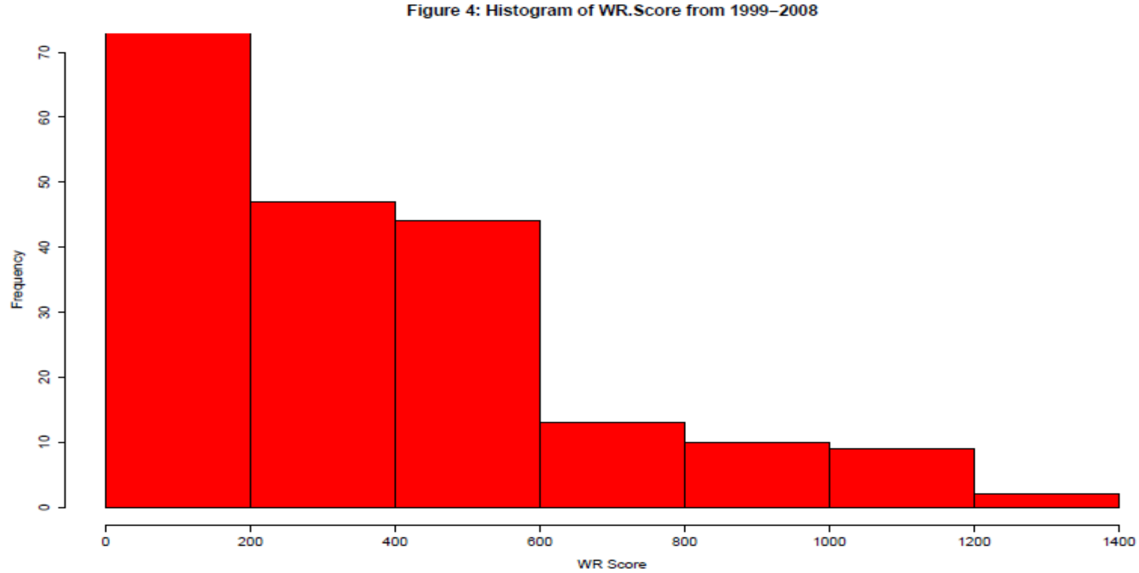
Looking at **Figure 4**, the mix of college and combine variables in this final model helped it receive a higher $R^2$ of 0.302 than the sum of the individual $R^2$ values found earlier. The 40-yard dash coefficient estimate has the smallest P-value out of all the coefficient estimates and is 227.5. This implies that, all else fixed, a decrease in a player’s 40-yard dash time by one-tenth of a second (0.1) will improve his expected draft slot by 22.75, which is almost one whole draft round! This provides further evidence to support the notion that the 40-yard dash time is seen as very important by talent evaluators. Two other variables that weren’t part of the tree regression model were the BCS dummy and BMI. We can infer that a receiver in a BCS school, all else constant, will have a higher expected draft ranking by 40 slots. This is practically significant because 40 slots is the difference between being a 1st round pick and a 2nd round pick! This presents evidence to support the notion that BCS receivers are favored in the draft, possibly due to increased media and television coverage, and/or better facilities and coaching. The model also suggests that an improvement in a player’s BMI by one point will lead to an improvement in a player’s expected draft rank by roughly 6 positions. Although BMI has a more subtle effect on draft position than does being a BCS receiver, it does suggest that NFL evaluators, on average, will favor receivers with more lean muscle mass.

## V. NFL Performance Prediction

In this analysis, I will examine whether the information NFL evaluators have on wide receivers prior to the draft can accurately predict their NFL performance. The response variable I will use as a proxy for NFL performance is the (*WR_score*). (*WR_score*) is my own wide receiver pseudo-scoring metric that calculates the annual "yards" a receiver is worth. The formula is: **WR_score = (Total NFL Receiving Yards + 18*Total NFL Receiving TD's) / (Years in League)**. There is much debate about how many yards an NFL touchdown is worth; many analysts agree that it is worth more than 10 yards, but Chase Stuart (pro-football-reference.com blogger) conducts a state-by-state analysis of historical NFL game play, and using results from David Romer's famous NFL dynamic programming paper, suggests that a touchdown is worth 18 yards, which is why I use that estimate in my formula.

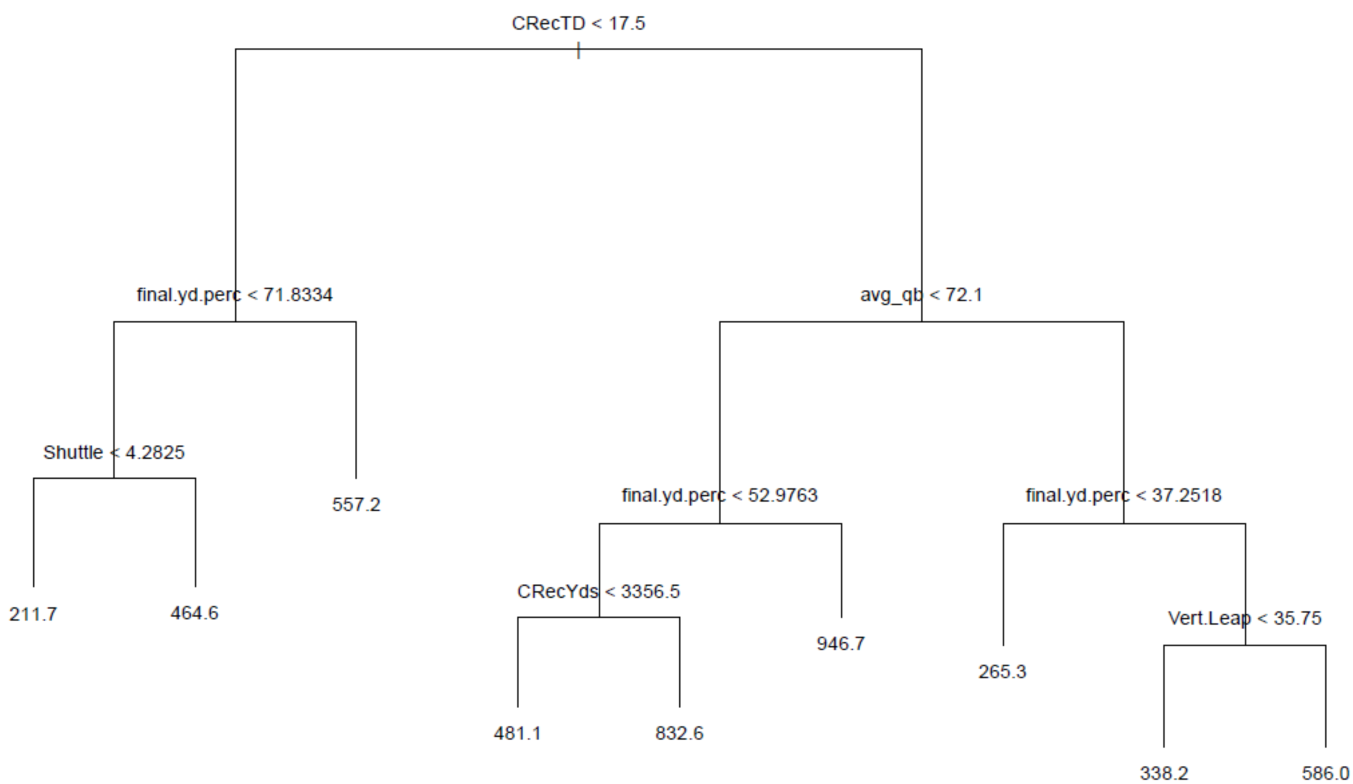
For this analysis, I kept the observations that had no missing values for the NFL and college statistics and had a (*WR_score*) greater than 25. I removed observations based on (*WR_score*) because I did not want the abundance of receivers who performed poorly to affect my prediction model for better performing receivers. Looking at **Figure 4** below, there is an evident right-skew in the distribution of (*WR_score*) because frankly there are more bad receivers than there are good receivers in the NFL. I will continue in my analysis with this (*WR_score*) cutoff of 25, although this is an area that warrants further research into. The predictor variables have remained the same from the draft pick analysis except for the inclusion of the (*avg_qb*) metric, which will help estimate the effect of a team's quarterback situation on wide receiver performance in the NFL.

**Figure 4: Histogram of WR.Score from 1999–2008**



As before, I will first discuss the results of my tree regression for (*WR_score*) and then examine the outcome from my linear regression model. **Figure 5** displays the tree diagram for this regression model predicting NFL performance. It is easily noticeable that the first variable split is on total college touchdowns (*CRecTD*), not total college receiving yards (*CRecYds*) as the previous tree regression model displayed. Although my previous results suggested draft evaluators tend to categorize receivers into "Over-achievers" and "Under-achievers" based on total college receiving yards (*CRecYds*), this tree model suggests that the bifurcation of the two groups based on total college touchdowns (*CRecTD*) is a split that more accurately predicts NFL wide receiver performance. These differing results could be explained by the fact that catching a touchdown in the end zone is a much harder job than catching a football in the open field due to the lack of field space in the end zone. Thus, receivers who are skilled at catching footballs in tight spaces in the end zone may have more success in the NFL because NFL defenders are much quicker than college defenders and usually can force receivers to catch footballs only in tightly spaced windows. NFL evaluators probably don't take this into consideration as they are judging different collegiate wide receiver prospects.

**Figure 5: Pruned Regression Tree – WR_score Prediction**



The "Under-achievers" portion of this tree (left side) provides similar insights to the insights obtained from the "Under-achievers" segment of the draft pick regression tree. This tree seems to suggest that receivers who didn't catch many touchdowns in college have to accumulate a high proportion of their total college receiving yards in their final season in order to be predicted of moderate NFL success. One crucial difference from the draft pick tree model is that the threshold for the final season's receiving yard percentage in this tree model (71.83%) is much higher than the threshold for the final season's touchdown percentage in the draft pick model (54.7%). This result suggests that a receiver needs a stellar final season to perform well in the NFL, and not just an above-average final season that the draft pick analysis seems to imply. Another noticeable omission in the tree is the 40-yard dash metric, which might point out that the 40-yard dash isn't as indicative of NFL performance as draft experts might think.

On the "Over-achievers" side of the tree (right side), the main division of the receivers is done by the team QB rating variable (*avg_qb*). The tree seems to predict higher NFL receiving performance when the team QB rating is below the 72.1 threshold, suggesting wide receivers do better on teams with bad quarterback situations. This outcome could potentially be due to the fact that teams with bad quarterback situations generally are poor on offense and that the team may have to rely on that particular receiver to generate offense downfield, which may spike his statistics but may not lead to team victories. Likewise, teams with excellent quarterbacks (like the Green Bay Packers and New England Patriots) generally do not only throw to one receiver and instead spread the football around all their wide receivers, to catch defenses off-guard.

The linear regression analysis introduces some new insights not already considered in the tree regression. In trying to determine the relative importance of collegiate statistics and combine data in predicting NFL performance, I estimated two OLS models, one regressing (*WR_score*) solely on the college statistics and the other solely on the combine results as I had done for the draft pick analysis. The former model obtained an $R^2$ value of 0.1457, while the latter model received an $R^2$ value of 0.02649. While it seems that collegiate statistics can explain relatively more variation of NFL performance than can combine statistics, both of their $R^2$ values are lower than their respective values from the earlier analysis, suggesting that there are other important factors that explain more of the variability of NFL performance that I am not taking into account. My final OLS model for predicting NFL performance for wide receivers is shown in **Figure 6**.

**Figure 6: WR_score Prediction – Final OLS Model**

Coefficients:

| | Estimate | Std. Error | t value | Pr(>\|t\|) | |
|---|---|---|---|---|---|
| (Intercept) | 824.9917 | 1301.1232 | 0.634 | 0.526812 | |
| CAvg | 133.9337 | 67.1756 | 1.994 | 0.047621 | * |
| CRecTD | 19.5449 | 5.9915 | 3.262 | 0.001313 | ** |
| final.yd.perc | 4.4497 | 1.3198 | 3.372 | 0.000907 | *** |
| X3Cone | -220.6536 | 101.5310 | -2.173 | 0.031011 | * |
| avg_qb | -36.2050 | 21.1161 | -1.715 | 0.088072 | . |
| I(CAvg^2) | -3.3991 | 1.9992 | -1.700 | 0.090741 | . |
| I(CRecTD^2) | -0.2260 | 0.1222 | -1.850 | 0.065929 | . |
| BMI | 36.7987 | 13.6856 | 2.689 | 0.007814 | ** |
| I(avg_qb^2) | 0.1993 | 0.1328 | 1.500 | 0.135194 | |

Adjusted R-squared: 0.216

The coefficient estimates with the two smallest P-values are the coefficients for Total College Receiving Touchdowns (*CRecTD*) and Final Season's Receiving Yard Percentage (*final.yd.perc*). That coincides with the tree diagram shown in **Figure 5**, as both variables were used as initial variable splits to maximize the decrease in sum of squared errors. As in my previous analysis, the BMI variable has a statistically significant coefficient estimate at the 1% level and seems to be positively associated with NFL performance. Thus, I can infer that wide receivers with more lean muscle mass tend to perform better, all else fixed, than receivers who are not as lean. Two variables that were practically significant in the draft pick prediction results, the 40-yard dash and the BCS dummy, are not present in either of the NFL performance prediction models. This suggests that although NFL talent evaluators pay close attention to the 40-yard dash time of a wide receiver and the type of school a receiver may come from, those factors aren't highly indicative of performance in professional football.

## VI. Conclusions

After thoroughly analyzing factors that may affect a wide receiver's draft position and NFL performance, I'd like to summarize my main findings from both the analyses and conclude by discussing some of the limitations that I faced in this project.

In the beginning of the report, I mentioned the anecdote about the Detroit Lions to make a general illustration that NFL teams are not able to accurately judge whether wide receivers will be successes or failures in professional football. I was able to analyze some of the important predictors of draft rankings using parametric/non-parametric models and compare and contrast those factors with some of the meaningful predictors of NFL receiving performance. In my opinion, NFL evaluators primarily tend to focus on NFL combine metrics and collegiate statistics when judging wide receivers in the draft so I focused on these two sources of data in my analyses.

By using a tree regression model, I was able to form two separate wide receiver segments for both the draft pick predictions and NFL performance predictions. I called these groups "Over-achievers" and "Under-achievers" because the variables splitting the two groups were Total College Receiving Yards (*CRecYds*) for the draft pick analysis and Total College Receiving Touchdowns (*CRecTD*) for the NFL performance analysis and both of these variables are a measure of collegiate performance. In general, I found that wide receivers in the "Over-achievers" segment tend to get drafted higher and perform better in the NFL over wide receivers in the "Under-achievers" group. I also argued that the result of (*CRecTD*) being the segmenting variable in the NFL performance analysis over (*CRecYds*) could be explained because it is harder to catch touchdowns in the narrow end zone than it is to accumulate receiving yards in the open football field. I maintained that being proficient in catching touchdowns in college may translate better to the NFL because NFL defenders are skilled enough to force receivers to catch footballs in very tight windows.

Other factors I found important in both the draft pick and NFL performance modeling were the 40-yard dash, the BCS dummy, and BMI. Both the tree regression and linear regression models indicated that the 40-yard dash time and the BCS-conference status were extremely important in the eyes of NFL evaluators when trying to determine the potential effectiveness of a college wide receiver. However, these same variables were not shown to be greatly predictive of actual NFL receiving performance, suggesting that the NFL draft evaluators shouldn't give these metrics a high weight in their analyses of wide receivers. It also suggests that talent evaluators should spend more time scouting the non-BCS wide receivers in order to "find diamonds in the rough." There was evidence to suggest that BMI was a significant predictor of both draft ranking and NFL performance. I also had information to suggest that wide receivers that get drafted to teams with bad quarterbacks tend to perform better in the NFL. I argued that those receivers drafted to bad offensive teams are relied upon to perform well and will get many footballs thrown his way, which can lead to artificially boosted receiving totals. On the contrary, teams with good quarterbacks tend to throw the football to many different receivers, reducing the performance statistics of any one particular receiver.

On the whole, college statistics seemed to be more predictive of draft rankings and NFL performance than did NFL combine data, suggesting that NFL evaluators do put more weight of their wide receiver rankings into the right areas (i.e. college performance, not combine). The $R^2$ values from both linear regression models gave some evidence to suggest this idea. However, the $R^2$ values were all less than 0.35, which implied that there is still a large amount of variation in NFL performance and draft rankings that I have not been able to account for in my analyses.

Because of the high amount of unexplained variability in my models, I lay out some limiting factors I encountered that constrained me in my analyses. My modeling did not take into account whether a receiver had injuries or not, which is important because a receiver's performance in college may be artificially dampened down due to injuries. I am also unable to quantify the character and motivation of a receiver; wide receivers that perform very well in college may shirk in the NFL and begin to perform poorly if they develop a bad work ethic. I was not able to acquire information on the types of offensive schemes that the wide receivers were a part of in college. If receivers played in offensive schemes that always throw the football then their collegiate statistics may be artificially inflated, whereas running-based offenses may artificially deflate a receivers' statistics. Additionally, my definition of wide receiver effectiveness in the NFL is only based on catching the football. Some teams draft wide receivers because they are exceptional at blocking defenders on running plays, an aspect of wide receiver play that I did not take into account. These factors are only a few of the prevalent variables that I have left out of my analysis, either due to lack of time in the semester or because some of them simply cannot be quantified. Regardless, I hope you find these insights I have described enlightening and hopefully I have been able to convince you that NFL teams, with their vast pool of resources devoted to preparing for the draft, still continue to flounder when it comes to drafting college wide receivers.