Path to the Podium: Analyzing Influential Factors in NFL Draft Player Selection

Maddie Weller and Jimmy Horvath

Professor Blunch

ECON 203: Econometrics

Introduction

Every spring, the National Football League hosts a draft for all thirty-two teams to select new players. Each team is allotted seven draft picks per year, one in each of the seven rounds. The first round of the draft attracts most fans' attention, as those players are the top prospects and can enhance a team's performance immediately. While there are thousands of college football players, merely thirty-two of them will get drafted in the first round. Our paper aims to identify what makes a player one of those thirty-two.

A first round draft pick has many implications, both for the professional team and the college team. For the professional team, this player will be good enough to immediately play a significant role in the following season and increase success. On the other hand, the college of the drafted player receives tremendous attention when it produces a first round pick. For example, the University of Alabama football team is well-known because many of its athletes are drafted in the first round. This likely is a motivating factor for many of the college's applicants and contributes significantly to the school's income through home ticket sales and revenues.

There are many variables influencing when a player is selected in the draft. Our paper will analyze these variables and answer the two research questions. First, what are the most important factors in producing first round draft picks in the NFL? Secondly, conditional that a player is a first round pick, what are the factors that affect where they're picked in the round? The results have implications for colleges as they will want to focus on the most influential variables so their football program can produce more first round picks and enhance the college overall.

Theory

$$P\,(1st\ round)$$

$$= f\,(W\%,\ EXP,\ R,\ H,\ C,\ HR,\ P,\ QB\ x\ R,\ OTOGC\ x\ W,\ Expenses\ x\ C) Pick\ Height\ in\ 1st\ Rd.$$

$$= f\,(W\%,\ EXP,\ R,\ H,\ C,\ HR,\ P,\ QB\ x\ R,\ OTOGC\ x\ W,\ Expenses\ x\ C)$$

1. WP (%) - A player's winning percentage is the number of wins in total games played in their final season of college football. NFL teams may consider this as a measure of a prospect's collegiate performance, which is a "a natural testing environment" to predict NFL performance (Williams, 2011). College football constants such as equipment and field size form realistic conditions to evaluate a player.
2. EXP – Football expenses represent the dollar amount spent by the institution on their football team. If a school spends more on their program, the players may have access to better coaching, phenomenal strength and conditioning facilities, etc. The team itself is likely to be more successful. An increase in spending of $1 million on a football program improved the probability of winning a game from 3.5% to 7% (McDonald, 2012). Thus, this variable is correlated with the above WP variable, and increases a draftee's resources, so we include it in the regression.
3. R – Black players are designated a value of 1, and non-black (mainly Caucasian and Hawaiian males) players were designated a value of 0. Black participants of the NFL combine from 2000-2018 were 8.1 percentage points more likely to be drafted (Gregory-Smith, 2023). Additionally, there have been less black than white draftees at positions such as QB, OLB, and TE (Gregory, Smith, 2023). On the other hand, the last white cornerback to regularly start in the NFL was Jason Sehorn in 2002. Thus, race will be an important variable to control for various factors like position.

4. H – height (in.) & W – weight (lbs.) - These variables vary greatly across positions. For example, running backs tend to be shorter players while tight ends are likely taller members of the team. Additionally, linemen tend to weigh more based on the demands of the position, whereas a wide receiver is typically leaner to sprint faster and jump higher. Depending on what position an NFL team would like to acquire in their first-round draft pick, these variables will respond accordingly.
5. C – If a player played in the Southeastern Conference, he was given a value of 1 and 0 otherwise. The SEC has had the most first round draft picks of any conference in the last 16 NFL drafts and had 7 of the first 15 picks of the 2022 NFL draft (SECsports.com). The SEC is clearly a dominant force in producing top NFL draft picks, so we include it in our regression.
6. HR – If a player grew up in the Deep South area of the United States (AL, FL, GA, LA, MS, SC) he was designated a value of 1 and 0 otherwise. Football in the South has historically been very popular, as in 1980, despite accounting for 31% of the U.S. population, 40% of major college football players came from the region (McConnell, 1963). McConnell also suggests many football players raised in the South stay in the region to play their college careers. Most of the Southeastern Conference (SEC) schools play in these states, as do their historically better programs in the conference. Thus, this variable has an important correlation with the Southeastern conference variable and may possibly affect the talent of the draftee.
7. P – position (broken down into several dummy variables)

The last piece of categorical data is a player's position. This is the most complicated dummy variable, given there are multiple subgroups. We group player positions based on the

areas the positions lie on the field, the similarities in functions of the position, and similarities based on the factors in our regression, providing some constants. While it was important to group positionally based on data, we also wanted to "step inside the shoes" of football executives and those who know the game well. Regressions are simply data driven from the Stata program, lacking the "human element" of knowing the game of football. While we're no experts, we're both fans of the game and know it well. Grouping positions based on our own football knowledge and research brought human intuition into a strictly data-driven analytical model, which we thought was necessary to get accurate results.

Our first assumption was to separate the offensive players and defensive players, given their functions and locations on the field are opposite each other. Starting with the offense, we initially grouped quarterbacks and running backs together because they are in the backfield of an offense and they're the primary runners of the football. The only other group that runs the football are wide receivers, and they rarely do. Additionally, wide receivers are not drafted because they can run the football well, especially not in the first round.

However, we agreed this methodology came with a host of limitations. First, while they both run the football, the quarterback is the only position on the field that throws the football. This is their primary function, which makes them valuable. Additionally, as we compiled our data, we found the percentage of running backs drafted in the second round compared to the first round to be higher than the same ratio for quarterbacks. We considered this detrimental to our grouping. Thus, we grouped quarterbacks alone, given their reputation of being invaluable (QB num).

Next, the logical decision was to add running backs to the positional group occupied by tight ends and receivers. They are the three players offensively eligible to catch passes thrown by

the quarterback. Most NFL executives and fans consider these “skilled positions”. Additionally, we considered the similar heights of receivers and tight ends to pair nicely.

However, tight ends are usually heavier than wide receivers and running backs by a large margin. Although most first-round tight ends are picked because they can block and are versatile catchers of the football, we pondered how their heaviness would impact the weight results of skilled positions. As a receiver or running back, weight is deemed less important than other metrics like height, speed, or abilities to catch the football, for example. We thought weight would be an important variable in our regression which would eventually help answer our two research questions. Since players can change their weight through nutrition and strength, this variable could yield meaningful results. Thus, we excluded tight ends from our second grouping, which is the skilled positions: running backs and wide receivers (RBWR num).

The remaining offensive variables are tight ends, guards, tackles, and centers. Although this group is known in football vernacular as the “offensive line,” the tight ends differ to the three other positions. Thus, our third group consists of offensive tackles, offensive guards, and centers (OTOGC num). These men block for the members of the backfield, especially the quarterback when he drops back to throw. We chose to exclude tight ends even though their positioning and responsibilities are like linemen. This is because tight ends can receive passes, which tackles, guards, and centers are banned from doing. Most NFL teams value their catching abilities above blocking. Furthermore, the weight of an average first round tight end is significantly less than that of first round tackles, guards, and centers. Accordingly, tight ends comprise the fourth grouping of the offense (TE num).

Next, we introduced defensive players to the model. Combining cornerbacks and safeties was most logical as they both play on the edges of the defense, defending the receivers and tight

ends from catching the football when it is thrown. Cornerbacks line up against wide receivers, and they are lined up the furthest wide from the ball before the play starts. Although safeties normally play in the middle of the field in terms of width, they are the furthest from the ball in terms of the length of the field. This is where the similarity lies – both positions are lined up far from the ball, albeit in different directions. Additionally, their speeds, weight, and height are incredibly similar. Many players drafted even played both safety and cornerback in college but are listed as one or the other come draft time. Hence, cornerbacks and safeties occupy the first group of the defense (CBS num).

Left to group were tackles, ends, and linebackers. Tackles and ends comprise the "defensive line", according to the football world. They line up closest to the quarterback and try and disrupt the backfield. They, like cornerbacks and safeties do, possess similar constants that make it easy to amalgamate them. Alternatively, linebackers serve many functions, far more than the tackles and ends. They are the captains of defense who play in the middle of the field. Linebackers must be fast enough to cover receivers but heavy enough to tackle incoming football runners. They comprise every function of the defense and are lined up in a different spot than every member of the defense, which is why they warrant a group of their own (LB num), while ends and tackles are grouped together (DT, DE num).

Offensively, our groups correspond to these variables: (QB num), (RB, WR num), (OT,OG,C num), and (TE num).

Defensively, our groups correspond to these variables: (CB, S num), (LB num), and (DT, DE num)

Literature Review

Many previous studies have regressed NFL draft outcome against similar variables, such as height, school winning percentage, the college program's expenses, etc. For example, Pitts and Evans studied drafted offensive players between 2002 and 2013. Expenses of the draftee's college program were significant at the 1% significance level, and specifically, an additional 1 million dollars for a running back's college team increased their draft position by 3 spots (Pitts, Evans, 2019). They also found that a standard deviation increases in the winning percentage of a wide receiver's college team improved draft position by 12, significant at the 10% level (Pitts, Evans, 2019). However, the same variable was not statistically significant for running backs even at the 10% level.

In a study of one-hundred-and-twenty-four NFL quarterbacks drafted between 1999 and 2008, height had a statistically significant effect on a quarterback's draft position (Berri, Simmons, 2011). Height also is statistically significant on running backs and tight ends' draft positions, but not for wide receivers (Pitts, Evans, 2019). In general, in both Pitt's and Berri's studies, the significance of a variable's impact depends on the position of the offensive player. While these existing studies focus on major offensive positions (RBs, QBs, WRs, TEs), our study extends to key defensive positions like defensive ends and tackles which are equally as important and address this shortcoming.

In addition to those numeric variables, important categorical variables have also shown statistical significance. For example, playing in the Football Bowl Subdivision (the highest level of college football) had a statistically significant impact on quarterback draft position (Berri, Simmons, 2011). The FBS is comprised of major Division 1 schools with successful football teams, so by playing against one another, players get better. FBS teams are eligible for The

College Football Playoff, which is a prestigious 4-team tournament to decide the national champion. A committee decides on the top four teams to qualify for this tournament, so teams and players that play in this are often the best of the best. In fact, players with CFP experience were selected more than an entire round earlier (Mulholland, Jensen, 2014). Furthermore, within the FBS, there is a group of conferences referred to as the "Power 5". These are the top conferences in college athletics, often producing the best players. Pitts and Evans report that playing in a Power 5 conference had significant effects on draft results at the 10% level.

While these three studies report meaningful effects of playing at a major D-1 college football program, a shortcoming is the lack of attention to the talent of football in the South. We believe that not only is a player's college performance vital for NFL success but also the football culture in which the player was raised. In the southeast, football is extremely popular as discussed in the theory section. Therefore, our paper addresses this limitation by regressing draft outcomes against the player's home region and playing in the Southeastern conference.

Data & Descriptive Analysis

Our data is composed of all first and second round NFL draft picks in the last eleven years (2013-2023). Given there is no major source of data for this specific topic, we created the database with values from NFL.com and ProFootballReference.com. The school's winning percentage, their football expenses, and the height & weight of a player make up our numerical data, while a player's race, school conference, home region, and football position make up our categorical data. These are the independent variables included in both regressions for the two dependent variables: round 1 probability and pick height. In our first regression, we consider both rounds of the past 10 NFL drafts, which is about 630 observations. In the second regression, there are around half as we only include first round data to see what determines a player's pick number in the draft. Below we discuss some issues underlying our data and noticeable descriptive statistics.

First, the winning percentage includes postseason games, which could be a lurking variable. The better teams in the country have more potential first rounders because their talent is better. Thus, they have a higher chance to qualify for a bowl game against another talented team, which would decrease one of those teams' winning percentage. The team of a player drafted in 2013-2023 won an average of 71% of their games, with a standard deviation of 19.6%. The minimum winning percentage is .083, held by the 2022 Northwestern Wildcats, who 2023 NFL Draft eleventh pick OT Peter Skoronski played for. The maximum winning percentage is 1, held by 2023 Georgia, 2021 Alabama, 2020 LSU, 2019 Clemson and 2018 UCF, where eighteen different first-round selections were made.

Additionally, in school expenses, some data on smaller schools like Tulsa or North Dakota State University, for example, isn't publicly available. Thus, we have missing values for

many smaller schools with significantly lower budgets. This skews the average of school football expenses higher than it ought to be. For example, the average school football expenses of the 2013-2023 first round selections are $50,400,000. When we include the second round of each draft for the first regression, this average decreases to $48,453,362. This decrease makes sense given more second round picks come from slightly smaller schools. Northern Illinois spent the least amount of money out of any school on football, as their budget was $8,153,654 in 2014. On the contrary, the University of Alabama has consistently spent the most money from 2013-2023, as their yearly budget is $78,521,050. The University of Alabama has produced at least one first-round selection every year from 2013-2023.

For weight, the average first round draft selection from 2013-2023 weighed roughly 248 pounds. Including both the first and second round picks, the average is slightly slower at 245 pounds. Given a standard deviation of 47 pounds, there is significant variation in the weights of players. This is expected given the variety of positions and the corresponding weights of each position. For example, linemen raise this average significantly, because that position is generally the heaviest, which has important implications for our conclusions.

For the binary variables, the minimum is 0 and the maximum is 1. For Race, the average is .828, meaning out of the 2013-2023 first round selections, at least 80% of the players are black. This demonstrates the athleticism of those players and how their physicality prepares them well for the demands of the NFL. In our data set, about a third of the 2013-2023 selections played in the SEC. This justifies the inclusion of this variable in the regression since such a substantial portion of players played in the conference. It proves the SEC continually prepares players for the NFL through the caliber of competitive teams that compete in it. For the region variable, it depends whether the second round of picks is included. Of just first round picks,

40.5% grew up in the deep south of the United States, whereas for both rounds, 33% are from the deep south. This decrease suggests that the players drafted in the second round tend more to come from areas not in the deep south.

For positions, the cornerback and safety position group accounted for the largest portion of first and second round picks at 21.5%. The offensive line and defensive line each accounted for around 17% of picks, respectively. Meanwhile, only 6% of picks were for quarterbacks. The portion of these positions will be important to note as we explain our study's conclusions.

Estimation Strategy

Our study asks two questions. First, what are the most essential factors in producing first-round draft picks in the NFL? Secondly, if a player is a first-round pick, what factors affect where they are picked?

We use the linear probability model to measure a player's likelihood of being picked in the first round to answer the first question. The coefficients will suggest the expected change in percentage points that a player is selected in the first round, given the conditional explanatory variable. Given how the linear probability model estimates, some of these coefficients will be negative despite probabilities traditionally being between 0 and 1. Although that is a weakness of the model, it is the most efficient way (we have learned) to analyze the factors that most produce first-round picks.

We use the traditional OLS estimator based on the given regressors in the theory section to answer the second question. It is essential to explain that a "higher" draft pick is desired. A player wants to be #1 overall versus #10. Therefore, some of the coefficients produced by this model will be negative. This means a player is drafted earlier and "higher," so they are closer to being the #1 pick, which is what we are looking to observe. The variables that yield the most negative OLS estimator will be the most important and relevant for answering the second question in our study.

Results

The table below includes the results of both regressions. P (first round) is the first regression to analyze what makes a player of first round draft caliber. The Pick Height column is the second regression showing the variables increasing the height in which a player is drafted.

Variable	P (1st round)	Pick Height
Winning %	0.190 (0.111)*	-1.546 (3.024)
School Football Expenses ($)	2.54e-09 (1.40e-09) *	-6.72e-08 (3.79e-08)*
Race (1=Black, 0=Non-black)	0.125 (0.064)**	0.649 (1.599)
Height (in.)	0.026 (0.012)**	-0.930 (0.325) ***
Conference (1 = Southeastern, 0 = other)	-0.027 (0.049)	-0.851 (1.27)
Hometown (1 = Deep South, 0 = other)	0.043 (0.047)	0.771 (1.18)
QB	0.884 (0.497)*	-7.012 (3.433) **
QB x Race	0.032 (0.177)	7.725 (0.042)**
RB, WR	0.493 (0.490)	-3.175 (3.84)
TE	0.504 (0.501)	3.675 (3.715)
OT, OG, C	0.900 (0.506)*	Omitted (multicollinearity)
Weight	-0.003 (0.001)***	0.003 (0.031)
(OT, OG, C) x Weight	0.005 (0.004)	-0.032 (0.105)
CB, S	0.518 (0.489)	-1.637 (3.96)
DT, DE	0.807 (0.498)	0.870 (1.98)
LB	0.620 (0.493)	0.791 (2.85)
p value < 0.10 *		

p value < 0.05 **
p value < 0.01 ***

While we included many regressors, we are concerned with a select number. These are the variables that college programs themselves can impact, allowing them to produce more first-round and higher-drafted players to benefit their program. Factors such as school expenses, conference, weight, hometown, and position can all be influenced by the college program from a recruiting and operations standpoint.

While Pitts and Evans (2019) found that the expenses of the draftee's college program were significant at the 1% significance level, our data yielded significance at the 10% level on round 1 likelihood. This variation could be because Pitts and Evans' study is limited to offensive players, while we considered all positions. Extending the scope of our study to include all position types was an element we thought needed to be improved in the existing literature and could be interesting to consider. However, all defensive position dummy variables yielded insignificant results in both our regressions. Therefore, if we also limited the study to only offensive players, we would have seen results like those of Pitts and Evans regarding expense significance.

Our study also included a regressor for players' hometowns, specifically if they were from the deep South of the U.S. While existing studies included the player's conference and playoff appearances as regressors, we believed that a player's success began earlier than their college program. Specifically, if raised in the deep South, where football is prevalent, they were exposed and trained from an early age. Despite this methodology and efforts to expand the literature, our regressor for the Deep South was insignificant in both regressions. Initially, the result was surprising, and we were even more shocked to find that playing in the Southeastern Conference

was also insignificant in both regressions. Berri and Simmons (2011) also considered the prestige of the college's football team, but more broadly considered if the team played in the Football Bowl Subdivision (FBS). This includes the SEC and major conferences like the Big Ten, ACC, etc. They found that playing in this subdivision of college program was significant on draft pick, which Pitt and Evans' regression also concluded. Therefore, while our specific analysis of the Southeastern conference proved insignificant, existing studies suggest that a college team's prestige is still relevant.

Among all regressors, we found weight and the positions of QB and lineman most significant and relevant to our study's purpose. According to our regression, the player's weight is significant at a 1% level. For the general regression, the coefficient is negative, reducing the likelihood of being taken in the first round. This finding suggests that a college team should carefully train and nurture these players while monitoring their weight for agile positions like quarterback, running back, and wide receiver. On the other hand, linemen need to be heavier because their job is to block the opposing team from tackling the quarterback. Therefore, we included an interaction term to analyze the effect of linemen gaining weight. Clinically, suppose a lineman can gain just twenty-three pounds (½ the standard deviation among the data). In that case, the coefficient for lineman (-0.72) plus the increment of the interaction term (Weight x lineman = 0.005) shows the player is 62% more likely to be taken in the first round. For pick, this interaction is not as significant. Nevertheless, this is a new insight that existing literature has not considered. Studies, such as Pitt and Evans (2019), consider height in various positions but not weight and lineman.

Our results also demonstrate the value of the lineman's role on an NFL team and the quarterback position. Being a lineman significantly (at 90% level) increases a player's first-round

chances, showing that NFL teams value these players. Given the previous analysis, our study shows the importance of this position and the relevance of added weight in increasing round 1 likelihood. In addition to the significance of the lineman position, our study also finds the quarterback incredibly significant. While this is widely accepted in the literature, our study provides support as the quarterback position significantly increased first-round likelihood (90% level) and draft height (95% level). To be precise, within the first round, a player in the quarterback position is drafted seven spots earlier. While this is important, our study yields an interesting finding regarding race and quarterback. In analyzing the interaction term of race (1 if black) and quarterback, a black quarterback is predicted to be drafted seven spots later than a non-black quarterback. This suggests a "penalty" for being a black quarterback, which is significant at a 95% level. While race is uncontrollable, and our study focuses on the factors college programs themselves can change, it is still a very thought-provoking finding of our study.

Our study supports findings regarding the significance of school expenses and the quarterback position on draft success. In efforts to broaden existing literature, however, we found no significance in our unique regressors of being from the deep South or being a player in a defensive position. Our study provides relevant findings regarding the importance of a player's weight, especially for linemen.

Conclusion

Our study provides convincing evidence that weight is vital in influencing whether a player is picked in the first round. While gaining weight reduces a player's likelihood overall, the interaction term we included for offensive linemen argues that weight improves that position's likelihood. Therefore, the college program should note the positions of their NFL prospects. For linemen, they should focus on the weight gain of those players through nutrition and weightlifting. On the other hand, weight gain hurts the likelihood for other positions, so the college should focus on fitness and appropriate weight maintenance for positions like QB, RB, etc. Given we are 90% confident that the quarterback position is more likely to be picked in the first round than the second round, and we are 95% confident quarterbacks are more likely to go seven slots earlier in the first round, college programs must focus on developing that position. Developing quarterback talent includes improving recruiting, coaching, film sessions, etc. Quarterbacks are often regarded as the "face" of the NFL team, as they are always the ones talked about on sports shows, news articles, etc. This exposure benefits the college, as discussed in the introduction. Therefore, our study suggests that one of the most effective ways to increase exposure is through the quarterback position. A college team should prioritize the quarterback for both their team's performance and the future benefits such a position can have.

As the NFL is a constantly evolving league, there are many exciting areas of future exploration. In recent years, many prominent universities and their athletic departments have switched the athletic conference in which the team participates in. For example, primary Pac-12 schools like Oregon and UCLA will leave for the Big Ten in 2024. In search of better competition, schools are switching conferences. While our study specifically considered the Southeastern conference for significant effects on draft success, it would be insightful to consider

other conferences like the Big Ten. As these conferences continue to grow and add more competitive teams, they will become more relevant in drafts.

Eventually, a different college football product will be shown on Saturdays to younger generations of football players across the country. While watching and playing college football is already standard in the Southeast, the current growth of football in the United States suggests that other regions are training their players as early and as intensively as they are in the Southeast. Therefore, an extension of our study could include other regions to explore this growth and see the other places that produce the next generation of NFL talent.

References

Gregory-Smith, I., Bryson, A. & Gomez, R. (2023). Discrimination in a Rank Order Contest: Evidence from the NFL Draft. J Econ Race Policy 6, 215–231. https://doi.org/10.1007/s41996-023-00118-y.

Mirabile, McDonald & Witte, Mark. (2012). Can schools buy success in college football? Coach compensation, expenditures and performance. College of Charleston. https://mpra.ub.uni-muenchen.de/40642/

McConnell, Harold. (1963). Southern Major College Football: Supply, Demand, and Migration of Players. University of North Carolina Press. https://doi.org/10.1353/sgo.1983.0006.

SEC Staff. (2022, April 26). SEC leads nation for 16th consecutive year in NFL draft. *Official Website of the Southeastern Conference.* https://www.secsports.com/article/33821910/each-sec-player-being-selected-2022-nfl-draft

Berri, D. J., & Simmons, R. (2011). Catching a draft: on the process of selecting quarterbacks in the National Football League amateur draft. Journal of Productivity Analysis, 35(1), 37–49. http://www.jstor.org/stable/23883795

Mulholland, J. & Jensen, S. (2014). Predicting the draft and career success of tight ends in the National Football League. Journal of Quantitative Analysis in Sports, 10(4), 381-396. https://doi.org/10.1515/jqas-2013-0134

Pitts, J. D., & Evans, B. (2019). Drafting for Success: How Good Are NFL Teams at Identifying Future Productivity at Offensive-Skill Positions in the Draft? The American Economist, 64(1), 102–122. https://www.jstor.org/stable/26725798

Round.do - Printed on 12/12/2023 11:37:45 AM

```
1   ********************************************************************
2   *                    Maddie Weller and Jimmy Horvath                *
3   *                                                                   *
4   *   Analyzing Influential Factors in NFL Draft Player Selection     *
5   *                                                                   *
6   *                              12/11/2023                           *
7   *                                                                   *
8   *                         ECON 203: Econometrics                    *
9   ********************************************************************
10
11  capture log close
12  clear all
13
14  * This do-file contains multiple regressions trying to answer two fundamental questions. First, what
    are the most important factors in producing first round draft picks in the NFL? Secondly,
    conditional that a player is a first round pick, what are the factors that affect where they're
    picked in the round?
15
16  *First, we import the sheet in our Excel that has both first-round picks and second-round picks, as
    we are observing probability of being picked in the first round.
17
18  import excel "C:\Users\horvathj25\Downloads\504.xlsx", sheet("BOTH ROUNDS") firstrow
19
20  * Before running any regressions, we had to convert string variables into numeric variables. These
    variables are listed below, and they're all binary variables in our data. Their definitions are
    listed in our paper.
21
22  destring Deepsouthnum QBnum RBWRnum TEnum OTOGCnum CBSnum DTDEnum LBnum, replace force
23
24  *Initially, we wanted to run all the variables we considered to be important in answering our first
    question. This doesn't include interaction terms because we hadn't considered those at the time of
    creating this regression.
25
26  regress Round SchoolsW SchoolsFootballExpenses Racenum Heightin Weight SECnum Deepsouthnum QBnum
    RBWRnum TEnum OTOGCnum CBSnum DTDEnum LBnum
27
28  *These four regressions below were created on curiosity. We wanted to see if the variables had any
    statstical significance themselves, not that it mattered towards our regression.
29
30  regress Round SchoolsFootballExpenses
31  regress Round Weight
32  regress Round SchoolsW SchoolsFootballExpenses Racenum Heightin Weight SECnum Deepsouthnum
33  regress Round QBxRace OTOGCXWeight ExpensesXSEC
34
35  *This regression answers our first regression question of what are the most important factors in
    producing first round draft picks in the NFL? We included the interaction terms in this regression
    to see the desired effects.
36
37  regress Round SchoolsW SchoolsFootballExpenses Racenum Heightin Weight SECnum Deepsouthnum QBnum
    RBWRnum TEnum OTOGCnum CBSnum DTDEnum LBnum QBxRace OTOGCXWeight ExpensesXSEC
38
39  *After the first question is answered, we must use different data to answer the second question of
    if a player is a first round pick, what are the factors that affect where they're picked in the
    round? Thus, we clear all to remove the "BOTH ROUNDS" sheet, given we're only looking at first round
    data. Next, we import the sheet in our Excel that has only first-round picks, as we are observing
    pick height in the first round.
40
41  clear all
42  import excel "C:\Users\horvathj25\Downloads\504.xlsx", sheet("first ") firstrow
43
44  *Initially, we wanted to do use the same mweethodology we used in answering question one. We wannted
    to run all the variables we considered to be important in answering our first question. This doesn't
    include interaction terms because we hadn't considered those at the time of creating this
```

Round.do - Printed on 12/12/2023 11:37:45 AM

```
    regression.
45
46  regress Pick SchoolsW SchoolsFootballExpenses Racenum Heightin Weight SECnum Deepsouthnum QBnum
    RBWRnum TEnum OTOGCnum CBSnum DTDEnum LBnum
47
48  *These four regressions below were created on curiosity. We wanted to see if the variables had any
    statstical significance themselves, not that it mattered towards our regression.
49
50  regress Pick SchoolsFootballExpenses
51  regress Pick Weight
52  regress Pick SchoolsW SchoolsFootballExpenses Racenum Heightin Weight SECnum Deepsouthnum
53  regress Pick QBxRace OTOGCXWeight ExpensesXSEC
54
55  *This regression answers our second regression question of if a player is a first round pick, what
    are the factors that affect where they're picked in the round? We included the interaction terms in
    this regression to see the desired effects.
56
57  regress Pick SchoolsW SchoolsFootballExpenses Racenum Heightin Weight SECnum Deepsouthnum QBnum
    RBWRnum TEnum OTOGCnum CBSnum DTDEnum LBnum QBxRace OTOGCXWeight ExpensesXSEC
58
59
60
61
```