Does Running Fast Lead to Thinking Fast? Evidence from the NFL Combine and NFL Draft

Ryan Dukarm

Haverford College

DOES RUNNING FAST LEAD TO THINKING FAST?

Abstract

This research aims to analyze the theory of thinking fast and thinking slow by studying how decision makers in the National Football League respond to high stakes and high visibility athletic testing results from the NFL Combine in their decisions in the NFL Draft. The athletic testing information gathered from the Combine is the last piece of information from a yearlong scouting process and provides an ideal testing ground for whether decision makers in one of the largest and highest stake industries in the world are overreacting to more recent information in their draft decisions. This analysis looks to tests four events that are commonly associated with being important to success at four offensive positions and whether they predict not only draft position, but if they predict subsequent NFL success as well. Our results indicate that all of the key events that are seen as common knowledge do predict draft position for their individual position and in general are properly valued by NFL teams. However, especially when it comes to running backs and offensive linemen, there is room for improvement to NFL teams' decision making process in properly valuing events that are less commonly assumed to be associated with success. This research provides results generalizable to other industries by suggesting there is often more than meets the eyes to predicting success while hiring.

Across a variety of industries there are copious amounts of available information for decision makers to consider in their hiring and decision making processes. Firms deciding on new employees to hire, who they should promote, or where they should locate headquarters, factories or offices, acquire a wealth of information from a variety of sources. How they decide to value this information, though, varies across firms and individuals. This research looks to examine whether one specific entity, the National Football League (henceforth written as NFL), shows evidence of decision makers thinking fast or thinking slow, two competing theories of decision making, by examining the relationship between the NFL Scouting Combine and the NFL Draft. Thinking fast is valuing the most recent information most strongly when making decisions, possibly ignoring older information during the decision making process. This would be seen as NFL executives drafting players in response to strong performances at the NFL Combine's athletic testing events, rather than the act of thinking slow and valuing the entire scouting process as a better predictor of success at the NFL level.

Due to the private nature of most hiring processes by firms (and similar decisions made by organizations such as college admission departments) there is little quantifiable data on what type of information goes into each decision. That is why the NFL Draft provides an ideal scenario in which to examine hiring processes. The nature of teams selecting the exclusive negotiating rights of players provides an implicit ranking of the value of players in the eyes of the teams, and the data from the NFL Combine is readily available. Furthermore, player performance data at the NFL level is also available to examine the effectiveness of these hiring decisions.

This research uses NFL Combine, Draft, and performance data to examine how NFL teams value athletic testing as a signal of ability, a process that takes place late in the overall

draft process. Are teams valuing the Combine in their decisions on which players to draft? In other words: are NFL Combine events significant predictors of NFL Draft position at individual positions? And if they are (or are not), should they be? Does the Combine predict NFL success for certain positions?

Institutional Background

The NFL is the collection of 32 pro football teams across the United States. The NFL Draft is the yearly selection of eligible players into the NFL by all 32 teams. Each team is given the rights to one single pick in each of the seven rounds. These picks are tradeable, both during the draft and in the weeks, months, and even years leading up to any draft. There are approximately 256 draft picks per draft, though there is fluctuation year to year on that number due to compensatory picks being awarded (additional picks in rounds three through seven for teams who lose too many players in free agency) or teams losing picks for breaking rules. In each round picks are given to teams in reverse order of standings, with the worst team in the NFL from the previous year selecting first in the following year's draft. Selecting a player in the draft gives a team exclusive rights to that player for the duration of their four year entry level contract. Players can refuse to sign their contract after being selected and be reentered in the pool of player's in the following years draft but this did not happen at all over the span of this dataset. All players who go undrafted are free to sign anywhere they please, and teams have a limited amount of cash they can pay to undrafted free agents.

Teams scout college football players during their regular season games all fall of the year prior to the draft in April/May. Many of the top players are often invited to a series of all-star games where they can practice and play under the eyes of NFL coaches, scouts, and executives.

Teams are allowed to host a limited number of prospects to their facility each year in the run up to the draft as well.

The NFL Combine is the last major event in the run up to the NFL Draft each year. This event is run by National Football Scouting Inc. and players attend the Combine by invitation only. National and BLESTO are the two main national scouting services that are employed by NFL teams in addition to their own scouting departments. These groups work to gather information to assist teams in their scouting process. Together, they represent a total of 28 of the 32 teams, per National's website, but all teams have a say in who is invited to the Combine itself. The directors of National and BLESTO, along with various representatives from NFL teams, form a committee every year to review and vote on which players should be invited. The goal of the Combine selection process is to invite all players who will be drafted in that year's player selection draft. There are always players selected who were not invited to the Combine, but overall this a strong grouping/collection of players. The Combine can accommodate up to 335 players each year, though the actual number of invitees varies slightly year to year.

In all but the most extreme circumstances, players invited will attend, as this event has massive career implications for prospects. As such, selection bias of a more talented than average player or bias towards lesser players who are desperate for exposure filling the places of talented players who chose to sit out (as is the concern for some All Star and college bowl games) is not a concern with this research. A key aspect of the Combine is comprehensive medical examinations, so top prospects who are injured are just as likely to be invited as those who are healthy. The Combine itself originated as a way to get universal medical exams on NFL Draft prospects. This is still a key part of the Combine, which now also offers teams the opportunity to

request 60 players for 15 minutes interviews, media availability, on-field position specific football drills, and comprehensive athletic testing.

The athletic testing aspect of the Combine provides the basis of data for this analysis, and the individual events, how they are collected, and what they assess are explained in the data section below.

Literature Review

This research connects to two strands of literature in addressing how decision makers in the NFL value the different events of the NFL Scouting Combine in relation to a collegiate prospects' NFL Draft potential. Copious literature has been devoted to different professional sports league drafts as ways of modeling hiring processes due to the large amount of available performance data from before and after the draft, as well the implicit ranking of the value of employees entering the market place by their draft selection position. This research looks to do the same by utilizing the environment controlled arena of the NFL Scouting Combine as a measure of ability that all firms (teams) have access to for all prospects who attend.

The first strand of literature this research continues to examine is how thinking fast and slow, developed from Kahneman's (2011) seminal work *Thinking, fast and slow* (Ichniowski & Preston, 2017; Kahneman, 2011), impacts high stakes employment decisions. Thinking fast would be making decisions while over-emphasizing the most recent information provided, in this case the NFL Combine results, rather than going through a slow methodical process of weighing the combine as one piece of a puzzle that has been developing for a long period of time, including actual college football games, playoffs, all-star games, interviews and private workouts. Thinking fast is the act of emphasizing more recent information more highly than older information. In the lens of this research, thinking fast would be an active reaction to the

NFL Combine results, and thinking too fast would be overvaluing the NFL combine events in the draft when these events do not predict additional production at the NFL level. If the Combine events predict NFL success, thinking fast should happen and NFL teams should respond appropriately on draft day. Thinking slow, on the other hand, would be not overreacting to better performance at the Combine if it doesn't predict NFL success. Thinking too slow would be a failure to react to events come Draft day if they do predict NFL success.

This research looks to examine the hiring processes of NFL teams to see whether teams think fast and react to strong combine numbers, even if they have not been good predictors of NFL success at a specific position, or whether teams think slow and value the combine accurately. This research most closely connects to Ichniowski & Preston (2017) who model thinking fast and slow by analyzing March Madness college basketball data and the NBA Draft to ascertain whether NBA decision makers thought fast or slow about unexpectedly strong performances by draft prospects in March Madness. Their findings suggest that front office executives in another high level, high visibility professional sports league thought fast and did react to strong performances by draft prospects in March Madness.

Not only did they think fast, though, Ichniowski and Preston also found that executives actually did not value strong unexpected March Madness performance enough and did show evidence of some thinking slow and valuing all information. In short, they thought fast but not fast enough. Decision makers were reacting to late games in predicting NBA success, but there was more predictive power to take advantage of there. Two players who were selected in the same spot of the draft, but one had experienced a "bump" in draft stock to get to this spot from a strong March Madness, performed differently at the NBA level, with those experiencing the bump outperforming their counterparts, all else equal.

With the NFL Scouting Combine providing a similar high profile event that is one of the last things that scouts and executives see of players before the draft itself, this provides another ideal laboratory to study thinking fast and slow. This research differs from Ichniowski and Preston as I do not analyze a change in draft stock from strong performance at the Combine as they do with NBA Draft prospects performing unexpectedly well in March Madness. This is largely due to the lack of availability of full 7-round mock drafts of the NFL Draft (the full length of the NFL Draft) in comparison to the 2 round NBA Draft mocks the others used (which is the full length of the NBA Draft). Instead, this research examines Combine events that are commonly seen as important to individual positions and whether they are examples of teams correctly valuing new information or overreacting to a player running fast in a controlled environment that does not adequately translate to the NFL level. Additionally, Ichniowski and Preston's research focuses on differences in the same variable (performance statistics in games) across different points of time by analyzing performance in March Madness compared to the regular season. Comparatively, the Combine is completely separate from college football performance and thus is a standalone set of variables separate from previous performance.

This research connects to a second set of literature that has used NFL Scouting Combine data to analyze NFL Draft tendencies, many of which have focused on the quarterback position (Berri & Simmons, 2011; Gill & Brajer, 2012; Pitts & Evans, 2018). Berri and Simmons' work is the pivotal paper in analyzing the NFL Draft by using measures of NFL production beyond simply games played and games started as some previous works did when using the Combine as a model for hiring practices (Hendricks, DeBrock, & Koenker, 2003; Kitchens, 2015; Massey & Thaler, 2012). While the previous papers laid a framework for analyzing the NFL Draft in this way, their choice to analyze all positions with the same dependent variables makes the results

somewhat less robust, as Berri and Simmons note. Using games started and games played as measures of the success of selections made by decision makers is a flawed analysis because which players play and start is another subjective decision being made to test the original subjective drafting decision (Berri & Simmons, 2011). Berri and Simmons narrowed their focus to quarterbacks, as did Pitts and Evans (2018) so they could use quarterback specific statistics as dependent variables of interest.

Pitts and Evans (2018) utilize OLS regressions of career NFL performance on Wonderlic test results, a mental aptitude test of 50 questions in 12 minutes given at the Combine, to assess the importance of cognitive ability in quarterbacks. Their research finds that across a multitude of dependent variables for career performance including games started, passing yards, wins, and approximate value (an all-encompassing integer value measure of ability based on individual performance statistics, whether the player went to the Pro Bowl, and some team performance measures), a higher Wonderlic score is a statistically significant predictor of increased NFL performance. Their research is part of a fairly split field among economists on whether or not the Wonderlic is a strong predictor of NFL success for quarterbacks. This split in the findings is largely a result of the differences in dependent variables and measures of NFL performance. My research will add to the existing literature by utilizing quarterback statistics that are more closely tied to NFL performance and ability rather than measures like yards or games started which are better measures of longevity than actual ability. Additionally, my research studies height as a key independent variable rather than Wonderlic score, which is not publicly available any more. Height is one of the most talked about attributes for quarterbacks in the run up to the draft and bears close study for its predictive power on Draft day and beyond.

Berri and Simmons (2011) developed NFL performance statistics on a per game and per play basis, including their own variables of net points and net wins produced by a QB per play, as well as an updated version of passer rating that better encompasses a QB's entire performance. Their work to develop per play measures of performance for a QB's career inspired much of the dependent variables used in this research. Berri and Simmons found that a variety of factors from the NFL Combine (height, weight, Wonderlic, and 40 yard dash time) predict draft position for quarterbacks, with taller, heavier, smarter and faster QBs being selected earlier. However, they found no significant results that any of these measures have a positive relationship on actual NFL performance. There are some concerns with their research, though, which is why quarterbacks will still be featured in this analysis. Namely, Berri and Simmons elected to cut off the eligible players for their dataset at those with 100 plays in a season rather than combining all their throws from their career into their analysis. This allows them only to analyze quarterbacks who are consistently playing in games, which is an incredibly small subset of NFL quarterbacks because only one quarterback plays at a time per team in all but the most unique situations. This analysis looks to analyze career performance to eliminate some of the selection bias present in season to season analysis.

This research looks to combine and extend elements of existing research to understand the causal relationship between various NFL Scouting Combine events, draft position, and NFL performance across different positions. Extending the innovation of Berri and Simmons (2011) to better measure NFL performance with advanced position specific statistics is an important way to make specific causal inferences about the NFL Combine. Further, this research looks to ascertain whether NFL teams think fast or slow, or some combination of the two in how they value the Combine in their decision making processes.

My research contributes to the existing literature on the relationship between the NFL Combine and NFL Draft in two ways. Firstly, this research looks at an expanded set of positions and athletic tests to better gauge the entirety of the scouting and draft process, rather than just focusing on one position. Secondly, I use more up to date and advanced measures of NFL performance that provide position specific measures of ability. Previous research that has used numerous positions in their research have kept the dependent variables of NFL performance very general, which hurts the precision of the results. This research uses position specific statistics as dependent variables that better represent the totality of a player's performance on the field.

Methodology

The empirical methods of this research follows strategies from previous literature by utilizing a set of linear regression models to estimate the effect of various combine events on the position of a player's draft selection and the effect of the same combine events on subsequent NFL performance. Regressions for positions of interest will be run separately to estimate the effect of individual events on draft stock and performance for specific positions. Specifically, this research looks to examine events that are commonly seen (by fans and the media) as important for various positions and finding 1) whether the NFL values those events and a strong or weak performance in those events impact draft position and 2) whether the NFL is correct in valuing (or not valuing) those events because of their causal impact on subsequent NFL performance.

The basic empirical model for analyzing the effect on draft position is as follows:

$$\mathbf{PICK_i} = \beta_0 + \beta_1 * \mathbf{Key\ Event_i} + \beta_2 * \mathbf{Other\ Combine\ Results_i} + \beta_3 * \mathbf{College\ Statistic_i} + \beta_4 * \mathbf{College\ Conference\ Dummies_i} + \beta_5 * \mathbf{Controls_i} + \varepsilon_i$$

Of importance to this research is that each position of analysis (quarterback, running back, wide receiver, and offensive line) are run in separate regressions. PICK is the dependent variable of interest, with lower numbers indicating earlier selection in the draft. The key event variable will thus change with each regression that is being run. The variable is meant to test the commonly held norm that a certain event is important for success as the position of interest. Those events/measurements are: height for quarterbacks, 40 yard dash for running backs and wide receivers, and bench press for offensive linemen.

Height is one of the most talked about attributes a quarterback has and tremendous media coverage goes into the official measurements at the Combine as the measurements listed on college websites are all unofficial and often vary greatly from the official height at the Combine. The common assumption about height for quarterbacks is that taller quarterbacks will do better because they need to be able to see over the much taller offensive linemen in front of them to make throws.

Both running back and wide receiver are positions that rely heavily on speed and overall athleticism, as they are the two main playmakers in space for football teams. As such, their pure speed is a highly talked about and highly televised spectacle during the Combine, as it is a simple number to interpret and react to.

Offensive linemen are the biggest players on the field and their main job is to block opposing players. Thus, strength is a huge focus in the evaluation of offensive line prospects. The bench press is also well covered and discussed in the media coverage of the Combine each year, as many people assume their performance on the bench press is an indication of how they'll be able to hold up against opposing players.

Key Event, then, is this independent variable of interest – this is the specific event that looks to test whether well reported and flashy events provide evidence of NFL teams thinking fast or slow in the drafting process. As such, β_1 is the coefficient of interest.

If there is a positive coefficient on the forty yard dash, three cone, or short shuttle variables, this indicates that a higher (slower) value for the test predicts a higher (later selection) value for the pick variable. Thus, slower players get picked later and faster players get picked earlier. This is evidence of thinking fast and reacting to the Combine.

Similarly, if there is a negative coefficient on the vertical, broad, or bench variables, this indicates that a higher number of inches jumped or bench press repetitions performed indicates a lower number for the pick variable. Thus, more explosive and stronger athletes are getting picked earlier as we would expect.

Other combine results include the results from all other athletic testing events, but also the height and weight measurements. For running back and offensive line, weight squared was also added when appropriate to the sample of data to study nonlinearity in the returns to more weight for the player.

College stats are the players' statistics from college measured in the same way as their later NFL performance statistics, explained in the data section below. Because of the nature of college football conferences with vastly different styles of play and levels of competition, a set of college conference dummies will be added to the model to help control for different levels of competition.

Controls is the vector of remaining controls, such as a draft year trend variable to control for changes in the era of football that may change the way certain positions, styles of play, or

production indicators are valued. This vector also controls for the class of the player entering the draft (junior or senior in college) with a dummy for senior year.

Additionally, as a modeling strategy to avoid losing players who elect not to run or cannot run certain events due to injury, players who do not participate in an event are coded with a zero rather than a missing value. Then, a series of dummy variables are created for each event for nonparticipation with a 1 for players who don't participate and a 0 for those who do participate in a given drill. This allows us to examine the effect of anyone who participated in any drills, not just those who elected to run in all of the possible drills. Most of these nonparticipation dummies are not included in the results tables but are merely used to maintain the sample size and precision of the model. ε_i is the random error term.

Turning to the other side of the research, NFL performance, the empirical strategy follows the same basic principles, though some of the control variables are adjusted.

$$\textbf{NFL Performance}_i = \beta_0 + \beta_1 * \textbf{PICK}_i + \beta_2 * \textbf{Key Event}_i + \beta_3 * \textbf{Other Combine Results}_i + \beta_4 * \textbf{Controls}_i + \varepsilon$$

The biggest difference is that the dependent variable is now a position specific measure of ability. Again, each regression is run separately for each position group. The dependent variable will be analyzed for rookie year performance, year three performance, and career performance on a statistics per season level of output.

The NFL performance measures are statistics meant to be all-encompassing measures that best represent overall NFL performance at each position. Full descriptions of the variables are given in the data section below.

Everything else in the model remains essentially the same as before, besides the introduction of the PICK variable. This allows us to make sure any driving influence from the

combine variable is not due to correlation with draft position, and all predictive power of the combine event is seen separate from the influence of draft position. The nonparticipation dummies remain to preserve sample size.

Controls is slightly adjusted in some specifications of the model now. When draft year significantly predicts draft pick position (which is the case for quarterbacks and running backs) a draft year and pick interaction term will be included in the NFL performance models to control for the changing valuation of the position across years of the draft. For other positions (wide receiver and offensive line) no draft year variable or interaction term are included as controls because the draft year is simply used as a control in the pick analysis.

These models allow us to examine differences in how the Combine effects both NFL draft position and later performance at the NFL level. Evidence of NFL teams overvaluing certain events and thinking fast without weighing all the available information would be significant results for the key event of interest in predicting draft position but then significant results that reverse the effect in the NFL performance regressions. For example, if the forty yard dash significantly predicts that faster players will get drafted higher in the first regression, but once we control for pick it predicts that faster players are actually less productive in the NFL, that would be evidence that NFL teams are thinking too fast and overvaluing the key event.

Conversely, evidence could exist that teams are thinking too slow and not adequately valuing results from a controlled and equal testing environment of all the top potential employees like the NFL Combine. Evidence for this conclusion would be finding significant results for the key event (or other specific combine events that may be more predictive of success) predicting NFL ability even while the did not predict draft position. Thus, decision makers would be wary of overreacting to athletic testing and are not adequately valuing information coming later in the

process that helps predict NFL success. They are afraid of thinking too fast and thus think too slow. Additionally, if there are results that the key event does have the expected effect on draft spot (faster, stronger, more explosive players are selected earlier) but the key event also significantly predicts NFL performance when controlling for draft position, that is indicative that while NFL teams are reacting to the information from the Combine, they are not reacting enough to take advantage of the potential gains available in NFL production indicated by better athletic testing.

Data

Data for this research come from a variety of sources to compile a comprehensive dataset of NFL draft prospects who were invited to and participated in the NFL Combine. The dataset includes Combine data and NFL performance data including: statistics, longevity of career, and positional data for individual players, the unit of analysis for this research. NFL performance data is gathered from Pro-Football-Reference which covers a variety of position specific measures of performance. This data is in player-season panel format, which allows for analysis by single year, by total career, or per season. The data spans the years 2001-2018 for all players who attended the combine between 2001 and 2015.

Combine data comes from the annual NFL Scouting Combine and is also collected though Pro-Football-Reference. It is gathered from the actual testing events at the combine itself that players can elect to run. The events and their subsequent scores for each of the players are explained below.

The *40 yard dash* is a straight line test of speed where players are timed at how fast they can run 40 yards from a standstill. The *short shuttle* (sometimes known as the 20 yard shuttle or the 5-10-5 drill) is test of quickness and agility that has players run 5 yards in one direction, turn

and run 10 yards in the opposite direction before running 5 yards to the starting point in the original direction. The *three cone drill* is a test of change of direction quickness. This involves running around a series of cones and making multiple 90 degree turns around them. The *vertical jump* is a measure of explosion where participants leap and touch measurements in the air to record their vertical jumping height. The *broad jump* is a standing broad jump where participants leap forward as far as possible from a stand still. Finally, the *bench press* is a measure of strength where participants bench press as many repetitions of 225 pounds as possible.

Results for the 40 yard dash, short shuttle, and three cone drill are reported in seconds. The vertical and broad jump results are reported in inches and the bench press results are reported in reps. Descriptive statistics for the combine results by position are presented in Table 1.

As can be seen in the table below there is a noticeable difference in times between offensive linemen and the other positions. In general, wide receivers and running backs tend to have similar results for many of the athletic testing events, with quarterbacks having generally worse results on average, followed by offensive linemen after that. The statistics are important to keep in mind as we move through the results of the analyses, as the coefficients for different positions in relation to both the magnitude of the mean and the spread of the standard deviation can often lead to differences in interpretation.

Table 1: *Summary Statistics for Combine Events and Measurements*

EVENT	Quarterback	Running Back	Wide Receiver	Offensive Line
40 Yard Dash	4.821	4.553	4.504	5.268
	(0.180)	(0.111)	(0.099)	(0.182)
	[4.33, 5.37]	[4.24, 4.93]	[4.22, 4.79]	[4.71, 6.05]
Vertical Jump	31.32	34.55	35.50	28.04
	(3.336)	(2.963)	(3.156)	(3.163)
	[21.5, 40.0]	[25.5, 43.0]	[26.0, 45.0]	[17.5, 37.5]
Broad Jump	110.2	117.5	120.2	101.0
	(6.657)	(5.620)	(5.468)	(6.625)
	[91, 128]	[103, 137]	[103, 139]	[74, 119]
Bench Press	19.64	19.85	14.94	25.51
	(4.152)	(4.377)	(4.214)	(5.164)
	[11, 26]	[8, 32]	[4, 27]	[9, 45]
Three Cone Drill	7.151	7.064	6.962	7.862
	(0.252)	(0.206)	(0.214)	(0.319)
	[6.55, 7.97]	[6.50, 7.71]	[6.42, 7.56]	[7.06, 9.12]
Short Shuttle	4.317	4.257	4.207	4.737
	(0.174)	(0.152)	(0.150)	(0.203)
	[3.87, 4.76]	[3.93, 4.73]	[3.73, 4.77]	[4.14, 5.56]
Height (inches)	74.74	70.45	72.67	76.57
	(1.679)	(1.774)	(2.256)	(1.525)
	[70, 79]	[66, 76]	[65, 78]	[72, 82]
Weight	223.1	214.3	201.6	313.1
	(12.06)	(13.92)	(15.18)	(13.42)
	[194, 288]	[168, 267]	[149, 241]	[263, 375]

Note: Mean values are presented for each event at each position. Standard deviations are presented in parentheses. Minimum and maximum values are presented in square brackets.

These Combine events represent the key independent variables of analysis. This dataset also includes variables of whether the player was drafted or not and at what specific selection. The draft selection spot will be the key dependent variable in analyzing which combine events are most important to NFL decision makers. Since many players who make it to the NFL go

undrafted (even though many were also at the combine) those players will be given a draft value of 260, which would be a few spots after the latest player was selected in any given year.

The aim of this research is to provide more thorough measures of performance than previous literature by utilizing more advanced and all-encompassing measures of performance at the individual position level. Previous literature (Hendricks et al., 2003; Kitchens, 2015; Massey & Thaler, 2012) has utilized games played and games started as dependent variables of interest across all positions. The issues with this approach have been discussed above and by Berri and Simmons (2011). This research instead looks to utilize statistics at the positions of interest that better encapsulate the actual play on the field of individual players.

For quarterbacks, passing yards are often seen as a simple and accurate measure of performance. However, because teams tend to throw more in games they are losing, simple passing yardage counts tend to overstate performance and do not take into account touchdowns, interceptions, and sacks – all of which are key parts of quarterback performance. As such, this research utilizes a slight transformation to Adjusted Net Yards per Attempt (ANY/A) as the dependent variable of interest. This statistic incorporates all the above elements of quarterback play into a per attempt measure of performance. Rather than a simple yards per attempt measurement, this statistic is calculated by assigning all non-yardage statistics (touchdowns, interceptions, and sacks) yardage values to add or subtract to the total passing yardage. The formula for ANY/A is below, courtesy of Inside the Pylon's football term glossary:

$$(\text{PASSING YARDS} - \text{SACK YARDAGE} + (20 \times \text{TOUCHDOWNS}) - (45 \times \text{INTERCEPTIONS})) / (\text{PASS ATTEMPTS} + \text{SACKS})$$

As can be seen above, touchdowns and interceptions are given multiplier values of 20 and 45 yards, respectively, and sack yardage is subtracted from passing yardage. In this analysis we will multiply the raw per season ANY/A by the number of attempts to get the total season or

career-long adjusted net yards. This helps control for games played and allows us to keep players who only throw a few passes each season but still gather statistics.

For running backs, adjusted rushing yards is preferred variable of interest. Similar to quarterbacks, running backs will receive a yardage adjustment for touchdowns equal to 20 yards for every touchdown scored and negative 45 for any fumbles either lost or recovered by the players own team. Though running backs score fewer touchdowns and commit fewer turnovers (interceptions/fumbles) than quarterbacks, they also gather fewer yards and fumbles recovered by either team should be seen as a negative. Thus, for ease of interpretation the level of the adjustments to total yards remain at 20 and 45 for touchdowns and fumbles, respectively.

For wide receivers, adjusted yards is also the variable of interest. The same touchdown/fumble adjustment that running backs receive will be applied to wide receivers as well.

For offensive linemen, who don’t generate traditional statistics, approximate value (AV) will be used as the dependent variable of analysis. Approximate value is meant to be an overarching statistic developed by Pro Football Reference to assign a single integer value to every player season to make cross position and cross time comparisons possible. While not perfect, AV provides a simple value for each player and gives players without their own statistics (especially offensive linemen) a numerical value we can use in the analysis. AV for an offensive linemen is a product of the team’s point production, their own starts, games played and whether they make a Pro Bowl or All-Pro team. This provides an explicit integer value for the ability of players based on objective performance statistics of overall team performance.

As can be seen in figure 1 below, the AV values for offensive linemen in single season or AV per season over the course of a career tend to be clustered at lower values. Player-season values for approximate value for all offensive linemen in the data set are presented below.

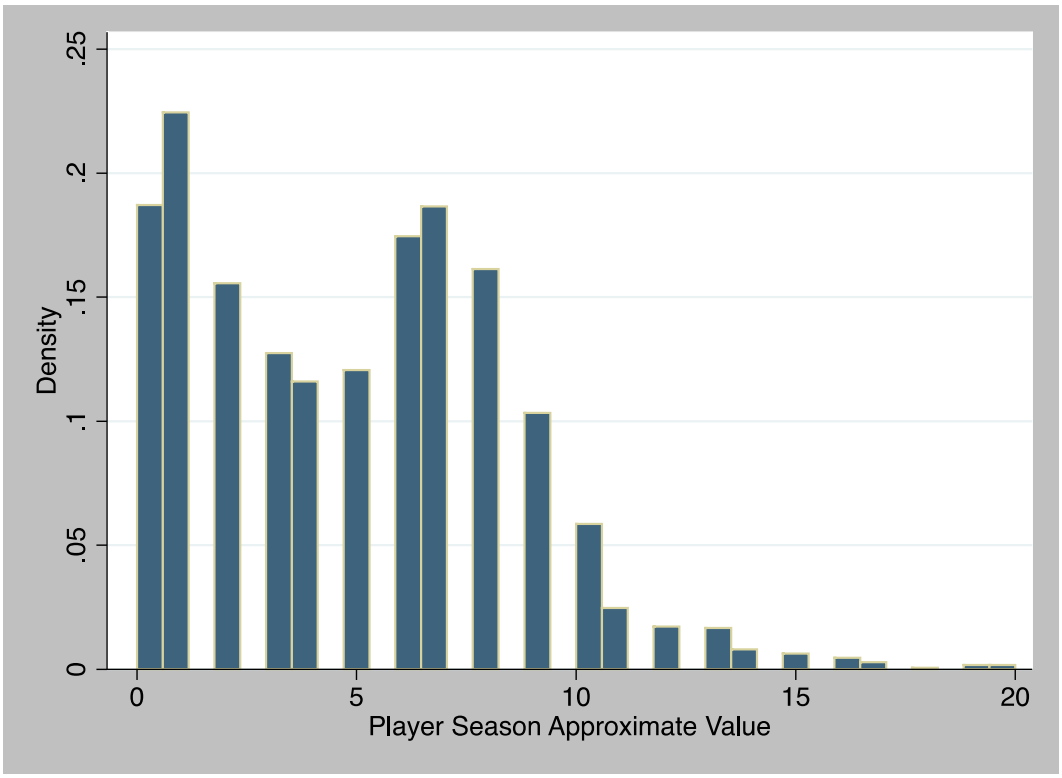

Figure 1. Histogram for the spread of approximate value for offensive linemen in each player-season in the dataset. This totaled 2,961 player-season scores of approximate value across all offensive linemen.

College football statistics are also gathered from sports-reference.com as a way to control for a players production in college as well as their games played and games started over their career. The same statistics from the NFL level will be collected from the college level as controls. However, sack data is not available for the college statistics, so quarterbacks will have adjusted yards (touchdowns/interceptions, but no sacks) controlled for rather than adjusted net yards. These stats have a high degree of correlation, and college adjusted yards should function

well as a stand in for college adjusted net yards. These data are for their final year of college performance, their class year, and the name of their school to add conference dummy variables. For offensive linemen, they gather no statistics and no universal resource for games played/started in college was available. For offensive linemen the only college control is their college conference and their class in their least season (junior or senior) which was gathered through the NFL's official website and various other draft tracking sources.

Results

Effect of Combine on Draft Position

In analyzing the analysis of the pick variable there is substantial evidence that the key event for each position does play a significant role in predicting draft position. Full results of the effect of the NFL Combine on draft position are presented in Table 2 below.

Quarterback results are presented in column 1. As expected, height predicts draft position in row 1 significantly to the tune of one additional inch of height for quarterbacks predicting a rise of approximately 16 spots in draft order, all else equal. With quarterback widely regarded as the most important position in football (Brooks, 2015; Feinstein, 2019), and substantial investments being made for quarterbacks in the draft (Mosqueda, 2018) a draft prospect's 'stock' can increase without any change in their previous college performance by simply measuring in at the NFL Combine taller than expected. The other significant predictors from the combine of a quarterback's draft position are weight (row 2) and forty yard dash (row 4). Overall, the NFL prefers taller, heavier, and faster quarterback prospects according to this study, which supports previous literature (Berri & Simmons, 2011).

Table 2: *Effect of Combine Results on Draft Pick by Position*

VARIABLES	(1) Quarterback Pick	(2) Running Back Pick	(3) Wide Receiver Pick	(4) Offensive Line Pick
1. Height	-16.40***	-2.608	3.011	-0.708
	(3.586)	(2.339)	(1.894)	(1.991)
2. Weight	-1.156**	-1.151***	-0.870***	-15.21**
	(0.567)	(0.313)	(0.290)	(6.850)
3. Weight Squared	X	X	X	0.021**
				(0.010)
4. Forty Yard Dash	106.8***	234.6***	324.9***	111.2***
	(39.58)	(35.70)	(30.51)	(18.80)
5. Vertical Jump	-1.768	-1.945	-2.223*	-1.600
	(2.295)	(1.527)	(1.242)	(1.247)
6. Bench Press	-9.603*	-0.796	-1.787*	-1.022*
	(5.176)	(0.879)	(1.017)	(0.615)
7. Broad Jump	-0.595	-2.041**	-1.025	-1.544**
	(1.147)	(0.929)	(0.760)	(0.652)
8. Three Cone Drill	39.09	59.55***	16.76	24.68*
	(32.58)	(20.62)	(20.26)	(14.45)
9. Short Shuttle	26.05	-27.30	-7.551	33.80
	(46.23)	(32.95)	(26.82)	(22.97)
10. Draft Year	2.789**	1.767**	1.248	-0.680
	(1.367)	(0.768)	(0.853)	(0.699)
Observations	275	414	637	816
R-squared	0.438	0.549	0.443	0.273

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

For running back prospects, as can be seen in column 2 row 4, the forty yard dash is a significant predictor of draft position. A decrease in forty yard dash time by one standard deviation (0.111 seconds) would help running backs' draft position by 26.21 spots. Faster players

at the running back position are clearly valued more by the NFL. Additionally, though, the NFL also values a number of other athletic and personal characteristics when drafting running backs. Both the broad jump (row 7) and three cone drill (row 8) are significant predictors of draft position with the expected signs of the coefficients indicating that better performances in either event leads to earlier selection. A one standard deviation increase in broad jump distance (5.620 inches) leads to being selected 11.47 spots earlier in the NFL draft, all else equal. Furthermore, the three cone drill also predicts draft position for running backs. A one standard deviation decrease (0.206 seconds) in the three cone drill time for running backs predicts selection 12.31 spots earlier in the draft. Running back is clearly a position for whom NFL evaluators are actively responding to tests of athleticism from the NFL Combine.

Interestingly, draft year is a significant predictor of NFL draft selection for running backs. As a result, a draft year and draft year times pick interaction term are included in the NFL performance models for both quarterback (which also had a significant coefficient on draft year) and running back to correctly control for draft year and draft pick in subsequent models analyzing NFL performance.

At the wide receiver position there are similar results to those of running backs for the effect of the forty yard dash on draft position, but the magnitude of the effect is larger at wide receiver. As can be seen in column 3 row 4, there is even more response from NFL decision makers to forty yard dash times than those from running back. A one standard deviation faster forty time (0.099 seconds) would predict a higher draft position of 32.21 selections.

The vertical jump (row 5) and bench press (row 6) are also significant, but at the 0.1 level for wide receivers. A one standard deviation increase in vertical jump height (3.156 inches) or number of bench press repetitions (4.214 reps) would lead to selection 7.02 or 7.53 spots earlier,

respectively, when holding all else constant. While only about a third of the sample of wide receivers participated in the bench press and the results should be taken cautiously, more than 500 of the 637 wide receivers in the sample jumped in the vertical jump which means the results should be taken a bit more to heart. Regardless, better performance in the forty yard dash, vertical jump, and bench press all predict higher selection in the draft for wide receivers. Weight (row 2) is also significant with a negative coefficient, indicating that heavier players will get selected earlier in the draft, all other athletic testing and measurements equal.

For offensive linemen and the bench press, the results also indicate that the common knowledge of what NFL teams value at certain positions holds up in the data. While bench press is not as overwhelmingly significant of a predictor for offensive linemen as other key events have been at other positions (row 6, p=0.097), it is a predictor of draft position. Full results for offensive linemen are presented in column 4 of table 1. A one standard deviation increase in the number of repetitions on the bench press (5.164 reps) would predict being selected 5.27 spots earlier in the draft. Again, this estimate does not quite compare to the magnitude of the other effects at other positions, but NFL teams are responding to an offensive lineman's bench press performance at the Draft.

Outside of the bench press, though, there are other athletic testing events that predict draft position for offensive linemen as well. The broad jump (row 7, p=0.018) and three cone drill (row 8, p=0.088) are both significant at the 0.1 level at least. A one standard deviation increase in broad jump distance (6.625 inches) would lead to selection 10.22 draft slots earlier, all else equal. Meanwhile, a one standard deviation decrease in three cone drill time (0.319 seconds) would predict selection 7.89 spots earlier on draft night, all else equal.

These results indicate that NFL teams are responding to multiple athletic tests for offensive linemen, but there are higher returns to the broad jump and three cone drill then there are to the bench press based on the magnitude of these coefficients in relation to each drill's standard deviation. Offensive line prospects training for the Combine who are looking to boost their draft stock should focus more on the broad jump then the bench press, as the magnitude of the effect of a one standard deviation change is near double in the broad jump compared to the bench.

Finally, offensive linemen are the one group of players in this study with non-linear returns to weight in terms of its effect on the pick variable. When testing additional models to see if any other positions required a weight squared quadratic adjustment, none came back significant except offensive linemen, which was the only position to thus have it included in the model. With a coefficient of -15.21 on weight (row 2) and 0.021 on weight squared (row 3), there are stronger returns to weighing in heavier up to a point, but NFL teams are turned away by additional weight at anything above 350.7 pounds.

The above results appear to strongly show that conventional wisdom is deeply entrenched in the NFL and football media landscape about which events are important to NFL success at four major offensive positions. All four key events show that NFL teams are aware and responding to this information and providing evidence that on some level, they are thinking fast and responding to information from the NFL Combine late in the draft process. Additionally, this analysis provides substantial evidence that while other events besides the key event at each position may not garner the same media attention, they do have additional power in predicting draft position.

Effect of Combine on NFL Performance

After assessing that NFL decision makers are actively responding to NFL Combine results, the next stage of analysis is to see whether these results adequately predict NFL performance or NFL teams are over (or under) reacting to the information gathered at the NFL combine. In general, significant results on any of the key events when controlling for draft pick position indicates a flawed decision making process on either end of the spectrum. If 'better' scores in the key event (taller, faster, stronger) significantly predict worse NFL production, this would indicate that NFL teams are thinking too fast, as the key event variable is undoing the predictive power of the event that will be captured within the pick variable. This would be substantial evidence of NFL teams thinking fast and overreacting to the information at the NFL Combine. Meanwhile, if there is significant predictive power of the key event on better NFL performance even while controlling for draft spot, then NFL decision makers are thinking *too* slow and should react even more to better performance in those key events, similar to the results of Preston and Ichniowski (2017).

Rookie year analysis.

To begin the analysis of NFL careers, we'll begin with analyzing rookie year performance in the NFL. This analysis will utilize all players who had any statistics in the same year as the year they were drafted. The dependent variables of analysis are adjusted net yards for quarterbacks, adjusted yards for running backs and wide receivers, and approximate value for offensive linemen. All dependent variables are in total rookie season stats. Full results for rookie year analysis across all positions are presented in Table 3 below.

Table 3: *Effect of Combine on Rookie Year Performance*

VARIABLES	(1) Quarterback Adjusted Net Yards	(2) Running Back Adjusted Yards	(3) Wide Receiver Adjusted Yards	(4) Offensive Line Approximate Value
1. Draft Pick	1,515**	-12.57	-1.759***	-0.0189***
	(651.2)	(131.1)	(0.239)	(0.00166)
2. Height	40.54	-11.86	11.84	-0.0164
	(69.56)	(16.48)	(11.08)	(0.1000)
3. Weight	-7.658	4.553**	1.260	-0.273
	(10.14)	(2.230)	(1.543)	(0.315)
4. Forty Yard Dash	26.34	-176.6	14.20	-0.961
	(639.5)	(216.9)	(200.8)	(0.994)
5. Vertical Jump	29.98	6.995	-10.32	0.0271
	(51.45)	(10.54)	(9.013)	(0.0633)
6. Bench Press	-267.0***	-4.664	-5.064	0.0112
	(64.27)	(5.566)	(7.874)	(0.0323)
7. Broad Jump	2.188	2.897	1.465	-0.0318
	(24.58)	(5.573)	(4.441)	(0.0353)
8. Three Cone Drill	211.9	-235.7	-29.86	1.637**
	(613.4)	(167.3)	(127.0)	(0.801)
9. Short Shuttle	390.7	501.1***	273.9	-0.786
	(892.7)	(188.4)	(179.7)	(1.213)
10. Draft Year	115.4***	2.152	X	X
	(37.38)	(12.05)		
11. Pick*Draft Year	-0.758**	0.00528	X	X
	(0.324)	(0.0653)		
Observations	95	238	297	419
R-squared	0.449	0.261	0.226	0.253

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Column 1 of Table 3 presents the results for quarterbacks. At the quarterback position, height appears to be valued correctly in the rookie year model, with no significant results for the height variable (row 1, p=0.562). There is no additional significant effect of height on total adjusted net yards for NFL rookie quarterbacks, which indicates it is properly valued already by its substantial effect on draft selection spot. Furthermore, the only combine event with any significant predictive power on NFL performance is the bench press, but only five of the 95 quarterbacks in the rookie data set performed the bench press so this can fairly easily be dismissed due to limited sample size.

In terms of rookie year performance, the data seems to show that a quarterback's height is properly valued on draft day.

For running backs, somewhat similar results hold in the rookie year analysis as presented in column 2 of Table 3. The forty yard dash is not significant (row 4, p=0.455) when controlling for where a player was picked. However, the short shuttle does have significant predictive power when it comes to predicting rookie year performance for running backs (row 9, p=0.006) but it actually predicts that faster short shuttle times lead to less productive running backs as rookies. A player with a short shuttle time that is one standard deviation faster (0.158 seconds) compared to an otherwise identical player would be predicted to run for 79.52 fewer adjusted yards in their rookie season, where touchdowns are again worth 20 yards and fumbles are negative 45 yards. The most likely explanation for this result is a combination of a few different factors. First, this is only one year of analysis for each player and comes from only their rookie year, when not all players get meaningful playing time, thus indicating sample size may be a bit of a problem. More importantly though, the short shuttle and three cone drill test very similar attributes of agility and change of direction. The correlation coefficient between the two results across all positions is

0.8782, indicating that these results should really be taken hand in hand with one another. As such, given that three cone drill significantly predicted draft position for running back (see Table 1, column 2, row 8) more likely than not this significant result for short shuttle with the opposite sign that would be expected is undoing the effect three cone drill had on pick, not indicating that less agile players will be more successful. And while this provides some evidence that NFL teams are thinking too fast at running back when it comes to the agility drills and rookie year performance, the lack of significant results at either the three cone drill or the short shuttle in either the year three analysis or the career analysis (in Tables 4 and 5, respectively) make this result a small tidbit and not much more.

At the wide receiver position, presented in column 3, nothing besides the draft pick variable is a significant predictor of rookie year performance. Similar to both quarterbacks and running backs, this does indicate that these tests are adequately valued by NFL decision makers when looking solely at rookie year performance. All events, and especially the key event of the forty yard dash (row 4), are adequately valued in the draft when forecasting rookie year performance.

Full results for offensive linemen are presented in column 4 of Table 3. The bench press (row 6) does not significantly predict anything beyond what is already captured by draft position for rookie year success as measured by approximate value. The three cone drill, though, is significant (row 8, p=0.042) and predicts that less agile offensive linemen will do better in the NFL. Similar to the results for running back, at least when it comes to rookie year performance, NFL decision makers appear to overreact to the three cone drill in the draft. Because the sign of the coefficient is suggesting that players who run the three cone drill in more time will have higher values of approximate value in their rookie year, while holding draft position constant, we

can interpret this to mean that NFL decision makers are thinking fast and overreacting to the three cone drill for offensive linemen, at least when it comes to rookie year performance.

Third year analysis.

Year three of an NFL career is typically where most players are expected to have reached their near final stage of development. There are plenty of late career success stories in the NFL, but as Dan Hatman, a former NFL scout and current director of the Scouting Academy, teaches his students looking to break into the game of football as scouts, many teams ask for three year projections from scouts to assess where that player will fit after a few years of NFL development. Thus, year three production is a good way to capture how a player has developed from potentially a set of raw athletic traits into a more complete and productive player.

Full results for year three analysis are presented in Table 4.

At the quarterback position, presented in column 1, we once again find no significant results for height (row 2) predicting any NFL success that is not already accounted for in the draft pick variable. Furthermore, no other events are significant predictors of production for quarterbacks in year three of their NFL careers.

For running backs, as can be seen in column 2 and row 4 of Table 4, there is no additional significant effect of the forty yard dash impacting year three performance. However, while vertical jump is significant at the 0.1 level (row 5, $p=0.066$), it actually predicts worse performance with higher vertical jumps. Much like the results for offensive linemen in their rookie year, the vertical and broad jumps are very similar tests as they both measure explosive athleticism. And while the vertical jump was not a significant predictor of draft position for running backs, the broad jump was (column 2 and row 7 of Table 2). Due to the high correlation between the two events (correlation coefficient of 0.8322), this result should be interpreted like

the results of the agility tests for rookie offensive linemen. Decision makers value explosiveness at the Combine for running backs, but it appears to be overvalued when it comes to predicting year three results.

At wide receiver, whose results are presented in column 3, similar results hold that no events are significant predictors of year three performance that is not already accounted for in the draft pick variable. NFL teams appear to properly value athletic testing when it comes to year three performance.

At offensive line, bench press is properly valued in draft position in regards to year three performance. Full results are available in column 4 of Table 4, but the bench press is insignificant when controlling for draft position (row 6, p=0.621). However, the forty yard dash is significant and predicts better performance with better forty times (row 4), indicating that NFL teams, even though they respond to faster offensive line forty times in the draft, are not valuing the forty yard dash enough. Since the mean weight of an offensive linemen in this dataset is 313.43 pounds, it makes sense that a test of speed is something that could easily be ignored by NFL decision makers. While this data does not contain the ten yard splits for the forty yard dash (the time it takes to run the very first ten yards – a test of short area quickness), that test may be more indicative of predicting performance for offensive linemen than the entire forty yard dash and is intriguing for future research. However, overall there is some evidence that NFL teams should respond to the forty yard dash more for offensive linemen, indicated both by their year three performance here and in column 4 of Table 5 in their career performance.

Table 4: *Effect of Combine on Year Three Performance*

VARIABLES	(1) Quarterback Adjusted Net Yards	(2) Running Back Adjusted Yards	(3) Wide Receiver Adjusted Yards	(4) Offensive Line Approximate Value
1. Draft Pick	372.0	-220.3	-2.107***	-0.0167***
	(667.1)	(149.5)	(0.291)	(0.00195)
2. Height	69.75	6.953	-4.427	0.00422
	(88.66)	(21.39)	(17.14)	(0.103)
3. Weight	3.563	2.810	3.499	-0.796*
	(11.63)	(2.700)	(2.201)	(0.425)
4. Forty Yard Dash	-600.6	-263.6	81.91	-2.093*
	(847.6)	(285.4)	(270.0)	(1.188)
5. Vertical Jump	-57.57	-29.68**	3.551	-0.00738
	(75.41)	(13.95)	(10.94)	(0.0733)
6. Bench Press	197.6	1.573	-12.36	-0.0176
	(253.2)	(7.268)	(8.469)	(0.0326)
7. Broad Jump	-18.56	9.246	-10.16	-0.0412
	(29.39)	(7.303)	(6.544)	(0.0439)
8. Three Cone Drill	-148.8	183.1	248.4	-0.391
	(945.5)	(179.5)	(162.2)	(0.818)
9. Short Shuttle	-414.9	-272.9	-237.2	-0.152
	(1,379)	(231.2)	(249.9)	(1.270)
10. Draft Year	32.49	-18.39	X	X
	(45.05)	(14.57)		
11. Pick*Draft Year	-0.189	0.109	X	X
	(0.332)	(0.0744)		
Observations	101	213	278	431
R-squared	0.325	0.274	0.200	0.211

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

These three year results seem to further indicate that on a limited basis of analysis in year three performance, NFL teams have a pretty strong measure on events that can predict success at four key offensive positions. While there are two examples of teams not valuing events properly (vertical jump being overvalued at running back and forty yard dash being undervalued at offensive line) the overall takeaway from the year three results is that there is mostly correct evaluation of the combine's athletic testing measurements.

Career analysis.

The most important piece of this analysis, however, is analyzing overall careers of all players in the dataset. The dependent variables are all in per season measurements created by simply summing the variable of interest (adjusted net yards for quarterbacks, adjusted yards for running backs and wide receivers, and approximate value for offensive linemen) and then dividing by total number of seasons. While a total measurement rather than per season measurement would likely be a more complete measurement of a player's achievement, due to recency of the data, many players are still playing, so this controls for ongoing careers and careers of different lengths. Full career performance results are presented in Table 5 below.

Beginning the career analysis with quarterbacks, which can be seen in column 1, we find no significant results that NFL teams are valuing various combine events incorrectly when it comes to the quarterback position. This result is intriguing because quarterback is the most important position in football, and without much argument, any major American sport. There are still plenty of factors that go into drafting quarterbacks that are not measured by this model, such as innate talent/skill, leadership, or developmental upside. However, when it comes to production in college and results from the NFL combine, it appears NFL teams have a good handle on predicting future success at the quarterback position from the data we have available. Because

Wonderlic IQ test scores are no longer available to the public they are not able to be studied in this research, but they remain a fascinating and intriguing piece of the puzzle when it comes to selecting quarterbacks in the NFL draft.

When looking at the career performance of running backs there is some evidence of NFL decision makers thinking too slowly and not properly valuing the forty yard dash. With a coefficient on the forty yard dash variable of -212.5 and decently strong statistical significance (row 5, p=0.084), these results indicate that even after the forty yard dash predicts draft position, running a faster forty *still* predicts additional NFL success to the tune of 23.74 additional adjusted yards per season for running backs who run the forty yard dash one standard deviation (0.111 seconds) faster. NFL decision makers are responding to the forty yard dash for running backs in their draft decisions as can be seen in Table 2, but as Table 5 shows, this response actually isn't enough. Much like the results of Preston and Ichniowski (2017), NFL decision makers should be thinking faster and reacting more than they are to the more recent information that is available at the NFL combine, at least when it comes to running backs and the forty yard dash. The forty yard dash is important and impacts draft position, but over the course of a running back's career it also predicts even more success.

Furthermore, weight (row 3) has a significant effect on NFL performance even when controlling for draft position. Heavier players are undervalued according to the model, as can be seen from the positive coefficient in Table 5 on the weight variable. Of course, for a position that requires speed and athleticism like running back, gaining weight above a certain threshold is a negative, thus the introduction of the quadratic element of weight squared (row 4). Combined with the negative coefficient on weight squared, heavier players tend to be more productive in

the NFL up until the threshold of 230.23 pounds, at which point additional weight predicts less production over the course of an NFL career.

For wide receivers, once again we find that the key event, forty yard dash, is being properly valued by NFL decision makers come draft time. Full results are presented in column 3 of Table 5. There is no significant effect of forty yard dash (row 5, p=0.949) that is not already picked up by the draft position variable. Weight is also a significant (row 3, p=0.082) predictor of success, much like at running back. Adding weight squared decreases any significance of weight into the model seeming to indicate that there are no negative returns to adding weight for wide receivers like there are for running backs, thus weight is the only variable kept. This is likely a result of the fact that on the upper ends of weight wide receivers can be converted to tight ends while there aren't similar position switches available for running backs. This seems to indicate that over the course of their career, heavier players will perform slightly better than NFL decision makers are currently accounting for in the draft.

Full results for offensive linemen are presented in column 4 of Table 5. At offensive line, the bench press is properly valued already in the context of the draft as there is no significant effect of the bench press (row 7) on approximate value per season when controlling for pick However, this analysis shows that there are multiple areas where NFL decision makers are not correctly valuing tests of athleticism for offensive linemen. Our first stage of analysis showed that there is significant response from NFL decision makers to better results in the forty yard dash, the broad jump, and the three cone drill when it comes to the NFL Draft (column 4 in Table 2, rows 4, 7, and 8, respectively). When analyzing the effect of all these events on career performance (Table 5 column 4), the forty yard dash should be even more strongly weighted in the selection of offensive linemen, as the negative coefficient seen in row 5 indicates that lower

(faster) times indicate a higher score in approximate value per season. Additionally, while the broad jump and three cone drill do not significantly predict any more success than the pick variable already does, the results in Table 5 show that NFL decision makers should be more responsive to the short shuttle (row 10) than they currently are in the draft. As discussed in previous results, the correlation between the agility drills likely indicates that given the results of short shuttle in the career analysis and three cone drill in the pick analysis, the takeaway from these results should be that the agility drills are both a factor in predicting NFL draft position and that they should be valued more heavily at the offensive line position in terms of consistent career production. A decrease in time of the short shuttle by one standard deviation (0.203 seconds) would result in an increase in approximate value of 0.326 per season. Given that the mean approximate value per season in the dataset for offensive linemen is 2.597 (with standard deviation of 2.849), this indicates the possibility of fairly substantial increases in offensive line production.

The fact that the offensive line seems to be the place where athletic tests may be overlooked by NFL decision makers isn’t altogether surprising given the nature of the position. Over 85 percent of the players in the data set weigh 300 pounds or more, so quick twitch athleticism is something that may be easy to overlook for a position that seemingly relies far more on size and strength. However, finding athletes that can still move efficiently and quickly even at such weights should be an asset to NFL team building strategies. This result specifically has interesting generalizable implications, that judgement upon first appearance may not always be the best indicator of what makes workers successful across different industries.

Table 5: *Effect of Combine on Career Performance Per Season*

VARIABLES	(1) Quarterback Adjusted Net Yards Per Season	(2) Running Back Adjusted Yards Per Season	(3) Wide Receiver Adjusted Yards Per Season	(4) Offensive Line Approximate Value Per Season
1. Draft Pick	237.3	-54.99	-2.091***	-0.0210***
	(270.4)	(67.07)	(0.129)	(0.000916)
2. Height	-6.059	-2.738	1.989	-0.0212
	(29.59)	(8.336)	(5.845)	(0.0507)
3. Weight	1.708	36.02**	1.567*	0.00973
	(3.757)	(17.21)	(0.940)	(0.00648)
4. Weight Squared	X	-0.0782*	X	X
		(0.0406)		
5. Forty Yard Dash	-102.1	-212.5*	-6.796	-1.023**
	(276.2)	(121.5)	(105.3)	(0.521)
6. Vertical Jump	10.45	-8.586	2.070	-0.0262
	(18.97)	(5.778)	(4.207)	(0.0331)
7. Bench Press	-34.15	1.437	-4.168	-0.0102
	(23.86)	(2.997)	(3.559)	(0.0166)
8. Broad Jump	-11.76	1.853	-2.898	0.0157
	(9.564)	(2.639)	(3.034)	(0.0178)
9. Three Cone Drill	-113.9	-36.13	28.41	0.628
	(264.9)	(92.53)	(61.93)	(0.382)
10. Short Shuttle	21.82	137.5	-29.51	-1.599**
	(415.2)	(114.1)	(97.88)	(0.632)
11. Draft Year	27.70	-3.432	X	X
	(30.39)	(7.882)		
12. Pick*Draft Year	-0.121	0.0264	X	X
	(0.135)	(0.0334)		
Observations	278	417	643	816
R-squared	0.439	0.436	0.389	0.481

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Robustness check.

To check for the robustness of these results, I also run the same regressions with approximate value per season as the dependent variable for quarterbacks, running backs, and wide receivers to see if there are similar findings to those with adjust net yards and adjusted yards. Full results are presented in Table 6 below.

Looking at the quarterback results in column 1 of Table 6, our earlier findings hold. No events produce significant effects on predicting approximate value per season for quarterbacks when controlling for draft selection. Overall we can feel confident in the results of career analysis that NFL teams are properly valuing the various combine tests for quarterbacks.

In analyzing running backs we find some variation from our earlier results. When using approximate value per season as the dependent variable no combine measurements are significant. While this does differ from the earlier results, the sign of the estimate for the forty yard dash, while not significant (row 4, $p=0.142$) is similar to that of the earlier results, as are the results of weight and weight squared (rows 2 and 3, respectively). While more overwhelming proof would have been more supportive, given the fact that approximate value is a bit restrained since it can only take integer values and is limited to fairly small numbers overall, these results lend at least some support to earlier results.

Table 6: *Robustness Check: Using Approximate Value Per Season for All Positions*

VARIABLES	(1) Quarterback Approximate Value Per Season	(2) Running Back Approximate Value Per Season	(3) Wide Receiver Approximate Value Per Season
1. Draft Pick	0.779	-0.0116	-0.0179***
	(1.026)	(0.652)	(0.00104)
2. Height	-0.102	-0.0194	-0.0541
	(0.120)	(0.0781)	(0.0495)
3. Weight	0.0139	0.256	0.0125
	(0.0153)	(0.162)	(0.00771)
4. Forty Yard Dash	-1.524	-1.606	-0.461
	(1.095)	(1.119)	(0.865)
5. Vertical Jump	0.0219	-0.0793	0.0380
	(0.0653)	(0.0545)	(0.0342)
6. Bench Press	-0.171	0.0120	-0.0381
	(0.104)	(0.0280)	(0.0292)
7. Broad Jump	-0.0335	0.00442	-0.0411
	(0.0361)	(0.0260)	(0.0254)
8. Three Cone Drill	-0.532	-0.762	-0.00320
	(0.965)	(0.853)	(0.508)
9. Short Shuttle	0.170	0.828	0.285
	(1.492)	(1.016)	(0.786)
10. Draft Year	0.0842	-0.00459	X
	(0.116)	(0.0771)	
11. Pick*Draft Year	-0.000399	-3.71e-06	X
	(0.000511)	(0.000324)	
Observations	278	417	643
R-squared	0.447	0.440	0.406

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

For wide receivers, the robustness check supports the earlier findings. Weight is the only predictor significant at 0.05 level in this specification and was the only predictor significant at the 0.1 level in the specification with adjusted yards as the dependent variable (Table 5). Heavier players tend to do better in the NFL and should be valued more strongly on draft day by NFL teams.

Conclusion and Discussion

The above results indicate that on the whole, the NFL is an example of an organization that is able to think slow and properly value large swaths of information when making critical hiring decisions. There are results that do indicate, however, that there are improvements that can be made in the NFL's decision making process and that athletic testing late in the information gathering process of the draft not only plays a role in where players at all four major positions of study get drafted, but also can impact performance in ways not currently accounted for in the drafting practices of NFL teams.

The main conclusions from the quarterback analysis is that the NFL has a good handle on what impacts success at this position, at least in terms of data that is publicly available. As mentioned above, now that Wonderlic scores are not publicly available, this area of study is not in the scope of this research but remains an intriguing way to study the helpfulness of intelligence tests like the SAT or ACT in future job performance across other industries.

The analysis of running backs is a bit more mixed and provides some evidence that there is room for improvement in the NFL's decision making process. Throughout the analysis presented above, various athletic tests presented significant results that there are flaws in the NFL's decision making process. While there are no overwhelmingly clear results across rookie, third year, and career-long dependent variables, there is evidence that various athletic tests as

well as weight are undervalued by the NFL for running backs. Running back is a position that relies more on pure athleticism than other positions, the player is simply given the ball and tries to run as far as possible. The results found above seem to indicate that NFL teams are not adequately taking into account the valuable information available to them from the NFL combine as it pertains to running backs. Forty yard dash predicts career success for running backs even more than is already being taken into account on draft day, as does weight up to the threshold of about 230 pounds. The NFL undervalues both faster and heavier running backs. While other events popped up as significant in various stages of the analysis, the consistent takeaways were fort yard dash and weight being undervalued.

At wide receiver, we find a different story for the same key event of interest. The forty yard dash and all other athletic tests are being properly valued by NFL teams in their decision making processes. Besides weight being narrowly significant in the career analysis, no other variables significantly predict NFL success besides draft position, which relies on the forty yard dash and college production more than anything else in the model.

Interestingly, weight seems to predict career success at both running back and wide receiver over the course of a player's career, but not in any of the individual seasons that are studied in this analysis. One of the most common talking points in the media leading up to the draft every year is that skinny players need to add bulk and become bigger to be able to handle the rigors and physical demands of an NFL schedule. Weight may not predict individual season performance, but it does appear to project overall longevity, durability and consistency on some level. This bears further research.

For offensive linemen, these results indicate that teams should be studying more deeply the events and information that is important to success at the NFL level. These may be the most

generalizable results of this study: often times there are gauges that can be better indicators of success than firms would think at first glance. Due to the sheer size of offensive linemen it's easy to overlook tests of athleticism and quickness as important to a position about dominating other big and strong players. However, the underlying data here seem to indicate that various athletic tests, such as the forty yard dash and the short shuttle, are undervalued come draft time for offensive linemen. Being able to move quickly as a massive player on the field is an asset teams need to more strongly prioritize in their hiring processes.

These results should provide evidence that NFL teams are able to think critically and analyze what leads to successful performance in the NFL. There are certain events that should be more widely valued by NFL teams at certain positions, especially in areas that common sense would lead us to believe don't matter for success at that position or for that type of player. These results should prove generalizable to a much wider set of industries, spurning research into what actually leads to successful professional performance beyond commonly held norms on what makes a successful employee.

References

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. https://doi.org/10.1007/s11123-009-0154-6

Brooks, B. (2015, July 27). Ranking each position's importance, from quarterback to returner [National Football League]. Retrieved from http://www.nfl.com/news/story/0ap3000000503855/article/ranking-each-positions-importance-from-quarterback-to-returner

Feinstein, J. (2019). *Quarterback: inside the most important position in the National Football League* (First Edition). New York: Doubleday.

Gill, A., & Brajer, V. (2012). Wonderlic, Race, and the NFL Draft. *Journal of Sports Economics*, *13*(6), 642–653. https://doi.org/10.1177/1527002511429575

Hendricks, W., DeBrock, L., & Koenker, R. (2003). Uncertainty, Hiring, and Subsequent Performance: The NFL Draft. *Journal of Labor Economics*, *21*(4), 857–886. https://doi.org/10.1086/377025

Ichniowski, C., & Preston, A. (2017). Does March Madness lead to irrational exuberance in the NBA draft? High-value employee selection decisions and decision-making bias. *Journal of Economic Behavior & Organization*, *142*, 105–119. https://doi.org/10.1016/j.jebo.2017.07.015

Kahneman, D. (2011). *Thinking, fast and slow* (1st ed.). Farrar, Straus and Giroux,.

Kitchens, C. T. (2015). ARE WINNERS PROMOTED TOO OFTEN? EVIDENCE FROM THE NFL DRAFT 1999-2012. *Economic Inquiry*, *53*(2), 1317–1330. https://doi.org/10.1111/ecin.12165

Massey, C., & Thaler, R. H. (2012). *THE LOSER'S CURSE: DECISION-MAKING & MARKET EFFICIENCY IN THE NATIONAL FOOTBALL LEAGUE DRAFT*. 35.

Mosqueda, J. (2018, December 15). NFL DRAFT: YOU HAVE TO MOVE UP FOR A QUARTERBACK [Optimum Scouting]. Retrieved from http://www.optimumscouting.com/news/nfl-draft-you-have-to-move-up-for-a-quarterback

Pitts, J. D., & Evans, B. (2018). 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. https://doi.org/10.1080/00036846.2017.1412081