April 18, 2025

Fourth Downs Decoded: A Predictive and Causal Analysis of Player Impact and Decision Making in the NFL

By

Simon Raymond

Supervisor & Examining Committee Member: Dr. Nicholas Manuel

A Thesis Submitted to Saint Mary's University
in Partial Fulfillment of the Requirements for
the Degree of Honours in Economics.
Halifax, Nova Scotia

Contents

Abstract

Fourth Downs Decoded: A Predictive and Causal Analysis of Player Impact and Decision Making in the NFL

By: Simon Raymond

Date of Submission: April 18, 2025

This research breaks down fourth downs in the NFL by identifying crucial player positions and evaluating how coaches incorporate different factors in their decision-making process. Using a newly engineered dataset that incorporated detailed player statistics from 2017-2023, we analyze the causal and predictive factors influencing fourth down decisions and outcomes. Our Generalized Inverse Mills Ratio reveals coaches factor in unobserved variables during fourth down decisions that can predict the likelihood of a fourth down conversion. However certain player archetypes, such as quarterbacks that excel in short passes, are not properly evaluated. We find significant impacts from other specific player positions, with offensive linemen showing strong joint significance on conversion but not on coaching decisions. The findings provide actionable insights for NFL coaches to identify key signals in fourth down situations and optimize player selection and team build for these critical plays.

Keywords

NFL, decision making, fourth down, player analysis, selection bias

1 Introduction

1.1 Context

The National Football League is a multi-billion dollar industry that has seen rapid growth within the American and international entertainment industry. This growth has placed more and more importance on the performance of NFL teams as they fight to win games to increase the popularity of their team as Davis and End (2010) argue that successful NFL franchises have measurable economic impacts on their local areas. This revelation has given rise to an age of data analytics in the NFL as teams seek to gain competitive advantages over their rivals. General managers and coaches then explore various strategic avenues.

In the NFL, a team's offensive possession is structured around a series of four "downs" where they attempt to advance the ball at least 10 yards to earn a new set of downs and maintain possession. Each down (1st down, 2nd down, 3rd down, and 4th down) is a chance to reach that 10 yard target that will give a team a new set of downs. If a team accomplishes this and recieves a new set of downs they return to being on "1st down" and begin their fight to reach another 10 yards. When a team reaches fourth down without achieving the needed yardage, they face a pivotal decision. Teams choose between three options: punt the ball to establish favorable field position for their defense, attempt a field goal if within reasonable range (worth 3 points), or "go for it" by running an offensive play to try and gain the required yardage. If a team attempts a fourth down and fails to convert to a new set of downs they give the ball directly to the opponent at the current field position. This field position makes it easier for the opposing team to score and therefore there is a incentive for teams to kick a field goal or punt the ball on fourth down. For many years in the NFL there seemed to be a consistent standard that when faced with a fourth down your team will kick a field goal or punt the ball for better field position. The only exception being the dying moments in a game when teams are desperate for a miracle.

In American Football many teams have shifted to being more aggressive on fourth down. Most famously are the Detroit Lions. Since the arrival of their current head coach Dan Campbell, the Lions adopted an aggressive strategy to match their aggressive "biting off knee caps" mentality (Birkett, 2023). While the Lions have seen success for the first time in years, they have also been criticized for their aggressive play calling. This was highlighted in the 2023-24 playoff divisional round game in which the Lions failed a fourth down attempt that was painted as unnecessary (Sporting News, 2024). After this mid-game failure there seemed to be a shift in momentum and the Lions lost the game.

1.2 Research Problem and Key Findings

A result like this causes one to ask the question "Did the Lions make the right call?". This question seems to be getting answered as "yes" by the current literature. However, we need to know if different teams should "go-for-it" or not "go-for-it" depending on their situation and team make-up. It could be argued that the Lions should have attempted the crucial fourth down in the 2023-24 playoff divisional round game. However, if the Panthers (which were a significantly worse team) were in that situation, it could again be argued that they should not have been as aggressive. The logic stems from how the Panthers would have a worse chance of being able to convert on fourth down due to a lack in player quality.

We must be wary of any recommendation that is given to a head coach. The truth is that we are not on the field, in the locker rooms, or in team meetings. This means coaches may know more then us in certain game time decision. We must approach this topic with the idea of being more practical and clear to coaches. If we are simply pointing our algorithms and giving black-box predictions we lose all credibility as one would be oblivious to the true depth of decision making in the NFL.

In our analysis we discover via our Generalized Inverse Mills Ratio that coaches are able to factor in variables that are unseen in our data and that are predictive of the likelihood of conversion success in their fourth down decision making process. However, certain archetypes of quarterback are not being properly evaluated and factored into decision making by coaches. Other positions such as offensive linemen, running backs, tight ends, and linebackers are also not being properly assessed by coaches on these “last chance” attempts.

1.3 Research Questions

This leads us to have a need to answer some key questions about fourth downs in the NFL. First, are coaches in the NFL actually better at making these decisions than analysts? Second, is there any predictive power in the performance of players? The type of data we have created and work with is incredibly detailed and the first of its kind. One would ask if practical applications of this method could be used in decision making. Finally, do these key variables about players have a causal effect on the outcome of fourth down attempts? Answering these questions will allow coaches to look for key signals in fourth down situations and to know which players to start on that fourth down if it is decided to attempt. This also can be applied in discovering specialty players that are overlooked due to poor performances in situations that are not similar to fourth down.

2 Literature Review

2.1 Coach's Decision Making

Much discourse revolves around the idea that NFL coaches are acting overly averse to risk, which is lowering their expected wins. Romer (2006) found that teams had begun to move towards a more conservative strategy in the NFL. He argues that teams value successful gambles more than the expected win percentage in a game. This is to say that a coach favours punting and kicking more then attempt a fourth down as they are easier to be seen as "successful". He theorizes that the poor decision making is either due to risk aversion or it is due to poor information. To further this point using matching analysis, Yam and Lopez (2018) quantified this conservative decision-making, finding that teams could gain approximately 0.4 wins per year by being more aggressive on fourth downs.

Goff and Locke (2019) found when revisiting Romer's framework that Romer's core findings are still held to be true. However, they argue that overly conservative calls are not due to poor decision making. Instead they point to risk aversion as they estimate that coaches are willing to give up two-thirds of an expected point to avoid the uncertainty of fourth down attempts. On top of this, there seems to be evidence that coaches are more cautious when their job is on the line. Owens and Roach (2018) found that in the NCAA coaches are relatively more conservative when they are more likely to be fired. At the same time they found when a coach was likely to be promoted they behave more aggressive then normal.

2.2 Momentum

If a team feels to be "on fire" should they be more aggressive since they feel they have momentum? A important area of literature is the fallacy of the "hot hand". The hot hand is a cognitive bias that leads people to believe that a person who has a successful outcome is more likely to have a successful outcome in future attempts. Gilovich et al. (1985) investigated the "hot hand" and "shooting streaks" in basketball. They found that both players and fans believed in the fallacy despite shots being independent of each other. Losak et al. (2023) similarly discovered that fantasy baseball users gravitated towards "hot" players. At the same time they were unable to identify a viable hot hand strategy in DraftKings DFS baseball.

Despite these common findings in other sports, there does seem to be some evidence of momentum existing in the NFL. Roebber et al.(2022, p. 2) defined momentum in the NFL as "the sustained increase in win probability by a single team over the course of at least 2 successive changes in possession". With this definition, they found that streaks of win probability in football are non-random and are in fact predictable with Artificial Neural Network Models. Lehman & Hahn (2013) looked to identify momentum across and within games in the NFL. Within-period momentum was found to encourage teams to take more risks. Negative within-period momentum was in turn found to encourage teams to take less risks. It was also discovered that across-period momentum has an effect only until a within-period momentum was established in a game. Due to its severity the fourth down is situation that contributes to the mental state of players and therefore deserves extra attention due to the potential to build or kill the momentum of a game.

2.3 Research Gaps

A gap in the current research is caused by the lack of quality data. Currently we observe many studies include team-aggregated grades or summary statistics about teams that are playing against each other. While some situations can allow this, our non-parametric models will be able to handle data with thousands of different variables. To take advantage of this, we will have information about every single player that is on the field when the ball is snapped. This will allow us to have better prediction power then previous researchers. These non-parametric models will also allow us to discover key player specific variables that can allow for further causal inference with specific on-field positions.

3 Data

3.1 Sources

Our data was pulled from two main sources. As a base the nflverse package provided in R gave us play-by-play data for the years of 2017 to 2023. This includes basic game information and the IDs of players that participated each play. The reason for our cutoff of 2017 is due to the NFL only putting tracking chips in players' jerseys as of the 2016 season. Since the 2016 season was the first year this was implemented there are many observations that lack the IDs of all 11 players on the field. For this reason we drop the 2016 season data.

The key part then is the merging of Pro Football Focus's aggregated weekly data. These statistics, which contain highly detailed information on individual players, are downloaded as premium player reports on a weekly basis and then aggregated to time lengths of 12 weeks. Players from the PFF data set were then merged into our base play by play data set based on a 12 step algorithm that matched different player IDs accross the two sources. The steps included pre-match ids, and then was followed by matches based on names, teams and positions etc.

Table 1: Key Data Summary

		Outcome Data							Select Data						
Variable	Available In	Mean	Median	SD	Min	Max	Zero Count	Zero %	Mean	Median	SD	Min	Max	Zero Count	Zero %
Conversion	Outcome Only	0.52	1.00	0.50	0.00	1.00	1923	47.85							
Attempt	Select Only								0.22	0.00	0.41	0.00	1.00	19970	77.98
Yards to Go	Both	4.12	2.00	4.58	1.00	34.00	0	0.00	7.77	7.00	5.73	1.00	46.00	0	0.00
Vegas Win Prob	Both	0.33	0.21	0.34	0.00	1.00	0	0.00	0.47	0.45	0.32	0.00	1.00	0	0.00
Spread Line	Both	1.98	3.00	6.35	-18.00	22.00	0	0.00	1.79	2.50	6.36	-18.00	22.00	44	0.17
Total Line	Both	45.43	45.50	4.55	30.00	63.50	0	0.00	45.01	45.00	4.56	30.00	63.50	0	0.00
Temperature	Both	60.84	70.00	15.35	6.00	93.00	0	0.00	61.73	70.00	15.14	6.00	97.00	0	0.00
Wind	Both	5.34	4.00	5.93	0.00	44.00	1604	39.91	5.42	4.00	5.73	0.00	44.00	9639	37.64
Def Stop Rate (Run)	Both	0.75	0.75	0.04	0.47	1.00	0	0.00	0.75	0.75	0.05	0.47	1.00	0	0.00
Def Stop Rate (Pass)	Both	0.67	0.67	0.04	0.39	0.86	0	0.00	0.67	0.67	0.04	0.37	0.91	0	0.00
QB Short Pass	Outcome Only	62.06	64.84	13.90	0.00	90.20	164	4.08							
QB Medium Pass	Outcome Only	63.36	66.97	17.08	0.00	99.10	209	5.20							
QB Deep Pass	Outcome Only	64.38	68.19	18.05	0.00	96.40	220	5.47							
Starter QB Short	Both	63.95	65.13	9.26	0.00	86.40	58	1.44	63.89	65.12	9.38	0.00	86.40	387	1.51
Starter QB Medium	Both	66.35	67.70	11.16	0.00	94.50	57	1.42	66.36	67.89	11.41	0.00	94.50	395	1.54
Starter QB Deep	Both	67.47	68.76	11.93	0.00	96.00	60	1.49	67.35	68.84	12.14	0.00	99.00	424	1.66

Dataset information: Outcome dataset: 4019 observations, 438 columns | Select dataset: 25608 observations, 123 columns

3.2 Variable Selection

Our variables that we work with fall into one of two categories. They are either statistics about players or about the situation of the play. The player variables consist of variables that describe who was on the field. We also have access to player variables of the starting players on each team based on depth chart data provided by nflverse. The depth charts are published before games by NFL teams. They state which players get general priority when choosing what player should be on the field. In both "on-field" and "starter" situations the players are sorted into columns based on depth chart position. For example, the starting QB (offense_player_1) is always in the starter_offense_player_1 slot to allow us to maintain consistency. Obviously our attempt models are forced to only use the starter player statistics as we cannot use the players on the field to predict the type of play that was called. This is because the types of positions not the field greatly change if a punt of field goal is called. For example, a quarterback is almost never on the field when a punt or field goal is executed. So in our data if a quarterback is found on the field we know that that play was a regular run or pass play.

4 Methodology

4.1 Tools

The following models include several categories of control variables: *Game Situation Variables* capturing tactical contexts like distance needed, timeouts remaining, and time factors; *Coach Variables* reflecting experience, tenure, background, and historical decision patterns; *Team Stats* measuring performance metrics, efficiency ratings, and formation tendencies; *Season FE* controlling for organizational factors unique to each team-season combination; *Coach FE* isolating individual coaching philosophies and tendencies; *Player Presence* binary indicators tracking which specific players or positions are on the field; *Control Def. Players* accounting for defensive quality, positioning, and scheme; and *Control Off. Players* measuring offensive personnel capabilities and skill ratings.

4.1.1 Predictive Models

After filtering out column with too many NA Values (at a threshold of 20 %) we maintain 50,000+ different player variables. This clearly is a situation where predictive tools such as Random Forest and XGBoosting thrive. In the use of XGBoost we first use boostrapping to tune for optimal hyperparameters. Second we will run the tuned model on 1000 train/test bootstrapped splits of data to report the Area Under the Receiving Operator Curve (AUC). This process will be done when predicting attempts and conversion. Similarly we can enact Random Forest to receive an Out-of-Bag AUC. Random Forest will also let us receive insight into what variables have high importance in predictions based on MDA and MDI. This ranking of variables not only allows us to feature select for XGBoosting but it also assisted in watching what variables could be considered for causal analysis.

4.1.2 Causal Models

For the causal analysis we will be employing a form of the Heckman correction. This is due to the selection bias in fourth downs. Since not all fourth downs are attempted, we do not have data on the fourth downs that never happened. We look to deal with this bias by using an evolved verison of Heckman's original model. In the original paper, it was proposed to use a two step process. In the selection step one would predict the probability of being selected while including a variable that is exogenous to the actual outcome that we are interested in. This estimated probability is then converted into the IMR and used to control for selection bias in the outcome step. This process however assumes linear relationships in the selection stage.

In our case we are in fact not dealing with linear relationships in the selection stage. This was discovered by how we have a increase of almost 20 %pts in AUC when predicting attempts with a XGBoost model instead of a linear model. Therefore our first step in correcting selection bias looks to predict the probability of an attempt of a fourth down. This prediction then is converted into a Generalized Inverse Mills Ratio that is not bound to the classical assumptions of Heckman's original IMR. We then place our GIMR in our outcome equation when estimating the conversion of a fourth down to control for selection bias.

4.2 Framework

4.2.1 Selection Model

Model the binary choice to attempt a play:

$$\text{Attempt}_i = \begin{cases} 1 & \text{if play is attempted} \\ \\ 0 & \text{if play is not attempted} \end{cases} \tag{1}$$

The probability is modeled as:

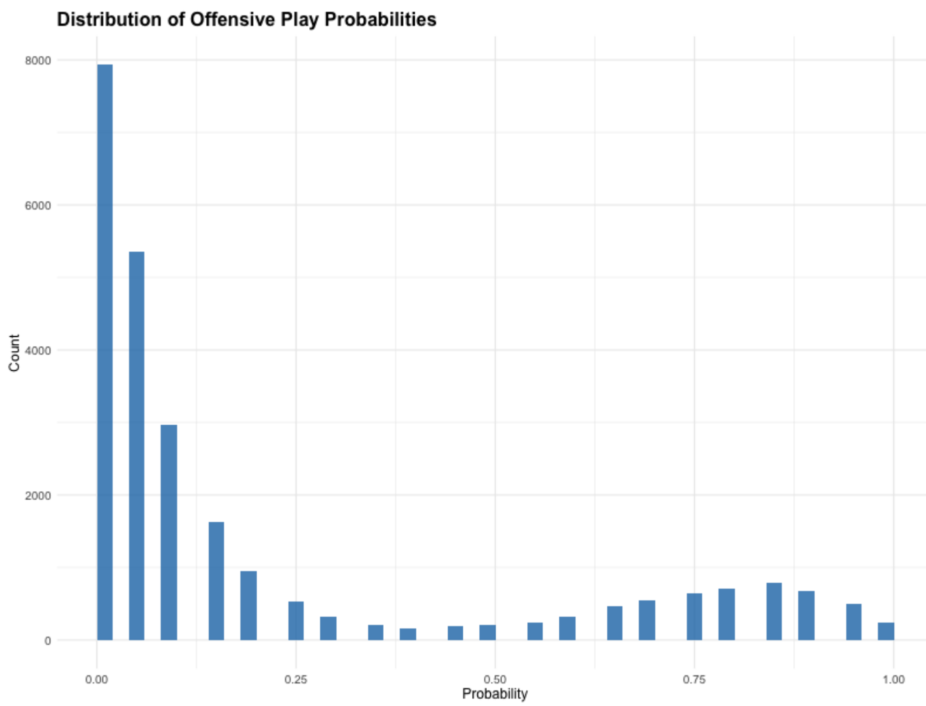
$$z_i = \Pr(\text{Attempt}_i = 1) = \text{RF}(\mathbf{X}_i, \mathbf{K}_i, \mathbf{P}_i) \tag{2}$$

where:

- z_i: predicted probability of attempt

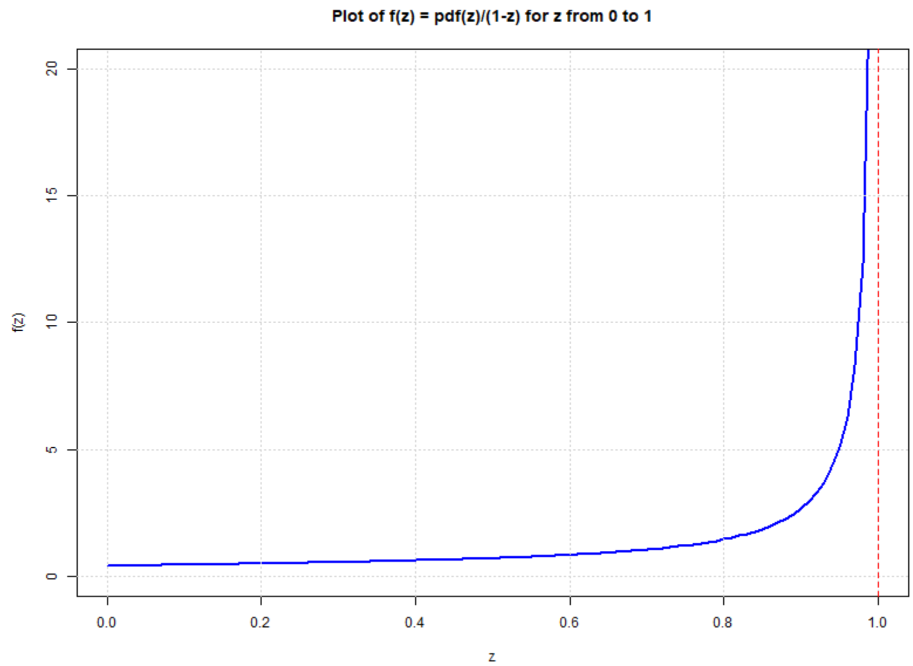
- $\mathbf{X}_i$: covariates excluding kicker and punter grades
- $\mathbf{K}_i$: kicker grades (exogenous)
- $\mathbf{P}_i$: punter grades (exogenous)

The distribution of z is:



The Generalized Inverse Mills ratio is:

$$\lambda_i = \frac{\phi(z_i)}{1 - z_i} \tag{3}$$



This functional form shifts the weights to emphasis situations where the probability of attempting a play is close to 1. We call this ratio "Generalized" because it is not bound to the classical assumptions of Heckman's original IMR. This is because the selection model does not contain latent variables and rather is non-parametric.

The first and second order conditions for λ_i are:

$$\frac{d\lambda}{dz} = \frac{\phi(z)(z^2 - z + 1)}{(1-z)^2} > 0$$
$$\frac{d^2\lambda}{dz^2} = \frac{\phi(z)[-z^3 + 2z^2 - 2z + 3]}{(1-z)^3} > 0$$

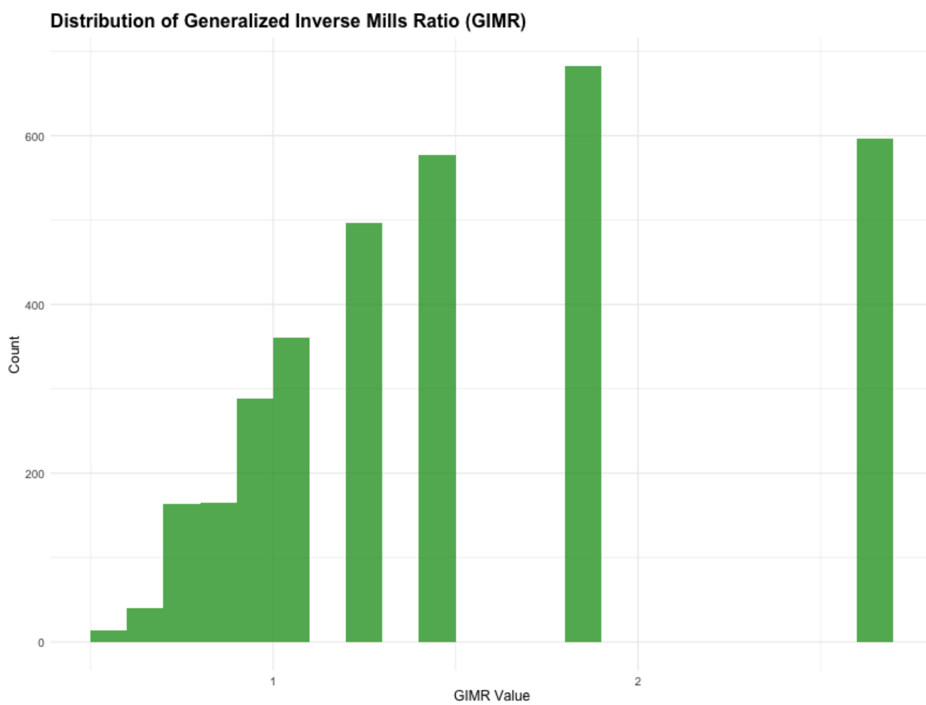
Thus λ_i is strictly increasing and convex in z_i.

4.2.2 Outcome Model

Model the binary conversion outcome for attempted plays:

$$\text{Convert}_i = \begin{cases} 1 & \text{if attempt is successful} \\ 0 & \text{if attempt fails} \end{cases} \tag{4}$$

The inputed GIMR then has the following distribution:



Linear probability model with selection correction:

$$P(\text{Convert}_i = 1|\mathbf{X}_i, \lambda_i) = \mathbf{X}_i\boldsymbol{\beta} + \theta\lambda_i + \varepsilon_i \tag{5}$$

where:

- $\mathbf{X}_i$: covariates excluding kicker and punter grades
- λ_i: generalized inverse Mills ratio
- $\boldsymbol{\beta}$: coefficient vector for main covariates
- θ: selection correction parameter
- ε_i: error term with HC1 robust standard errors

Note: Kicker grades ($\mathbf{K}_i$) and punter grades ($\mathbf{P}_i$) serve as exogenous variables in the selection equation but are excluded from the outcome equation for identification.

Models are estimated separately for offensive outcomes (grades, yards, completions) and defensive outcomes (stops, grades, tackles). A defensive values are of the team that does not have possesion of the ball when offensive values are of the team that does have possesion of the ball.

5 Results

5.1 Exogeneity of Kicker and Punter Grades

A key feature in our sample selection correction model, is the exogeneity of a variable that influences the selection equation. In American Football there is the luxery of using third down conversions as a selection bias free area. This is due to how conditions such as desperation or play calling cause a third down to be treated similarly to a fourth down by coaches. A fourth down is a teams last attempt for that offensive drive. While not in all, there are many cases of third down provide the similar conditions in terms of the coaches decision making. For example if the ball is on a teams own 20 yard line and they are on third and 10, the coach will treat this attempt as a “last chance”. This is due to the fact that if the team does not convert on third down they will be forced to punt the ball. There is no situation here where the team can even consider kicking a field goal.

We can first note that the grades of a teams kicker is positively significant in its effect on the decision to attempt a fourth down. Exogenity of the kicker is now required in the case of the outcome of fourth downs. When the kicker is not on the field during a fourth down they are unable to effect the decision making of the coaches or plays of the coaches, as there are no further decisions to be made. In a Third down situation we find that kicker grades do have positive significance on the conversion of a third down. This significance, is solely found in the middle of the field, where the kicker is more likely to be used if the third down is not converted. In the situations on third down that would mirror a fourth down, there is no statistical significance. When the kicker is not a part of the decision making process on third down, similarly to fourth down, it does not have an effect on third down conversion.

The following table 2 displays the effect of kicker and punter grades on the conversion of a third down in different ranges of area on the field. While the kicker and punters are not on the field they still show signs of influencing decision making when and only when they have a possibility to effect the future plays. For example we see that kicker grades are positively significant in the middle of the field. This is due to how if a team is more trusting of their kicker they do not need to keep in mind as much gaining an extra few yards to make the field goal attempt easier. They instead can solely focus on converting to a new set of downs and not having to fear a missed kick if they fail to convert. It is important to note a key part of NFL defenses in this process as well. In a situation that is third down and 10 yards to go, the defense is not trying to stop the offense from gaining 5 yards. Rather they position themselves to stop the offense from gaining 10 yards. This forces coaches to ask themselves "do I trust my kicker enough to run a play that has a higher chance of gaining 0 yards but also a higher chance of gaining the yards needed to convert". A similar logic applies to the punter grades as well. When a team is more comfortable with punting the ball they will call plays more conservatively. This over-conservative play reduces the chance of a conversion.

Table 2: Kicker Offensive Grades

Variable	Own 1-10	Own 11-20	Own 21-30	Own 31-40	Own 41-50	Opp. 49-40	Opp. 39-30	Opp. 29-20	Opp. 19-10	Opp. 9-1
LPM Model										
Punter Grades (12w)	-2.402**	-1.091	-2.110**	-4.109***	-0.862	-1.932*	-0.303	-0.193	-0.364	0.210
Kicker FG Grades (12w)	2.742***	-0.328	0.819	3.303***	2.120**	2.725***	1.650.	0.071	0.355	0.317
Probit Model										
Punter Grades (12w)	-2.514**	-0.986	-2.065**	-4.006***	-0.829	-1.852*	-0.429	0.064	-0.236	0.169
Kicker FG Grades (12w)	1.681*	-0.521	0.808	3.182***	2.058**	2.522**	1.582.	0.014	0.153	0.314
Logit Model										
Punter Grades (12w)	-2.376**	-0.998	-2.121**	-3.937***	-0.788	-1.797*	-0.406	0.011	-0.315	0.184
Kicker FG Grades (12w)	2.080**	-0.508	0.876	3.162***	2.084**	2.510**	1.595.	0.014	0.209	0.320
Controls										
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Player Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Player Presence	No	No	No	No	No	No	No	No	No	No
Season FE	No	No	No	No	No	No	No	No	No	No
Coach FE	No	No	No	No	No	No	No	No	No	No
Sample Size	988	3,023	6,865	6,924	5,805	4,945	4,271	3,897	3,478	3,080

[1] Values shown are t-statistics from respective model regressions with conversion as the dependent variable.
[2] Significance codes: *** p<0.01, ** p<0.05, * p<0.1, . p<0.15.
[3] Each column represents a 10-yard section of the football field, moving left to right from your own goal line (Own 1-10) toward midfield and then to the opponent's goal line (Opp. 9-1).

This significance of the kicker and punter grades in key areas of the field is due to the coaches being able to properly adjust their decision making for whether or not to attempt a fourth down. This is to say that coaches are properly recognizing how good or bad both their punters and kickers are and making adjustments on their playcalling.

As a note while punters do not show signs of significantly affecting the fourth down

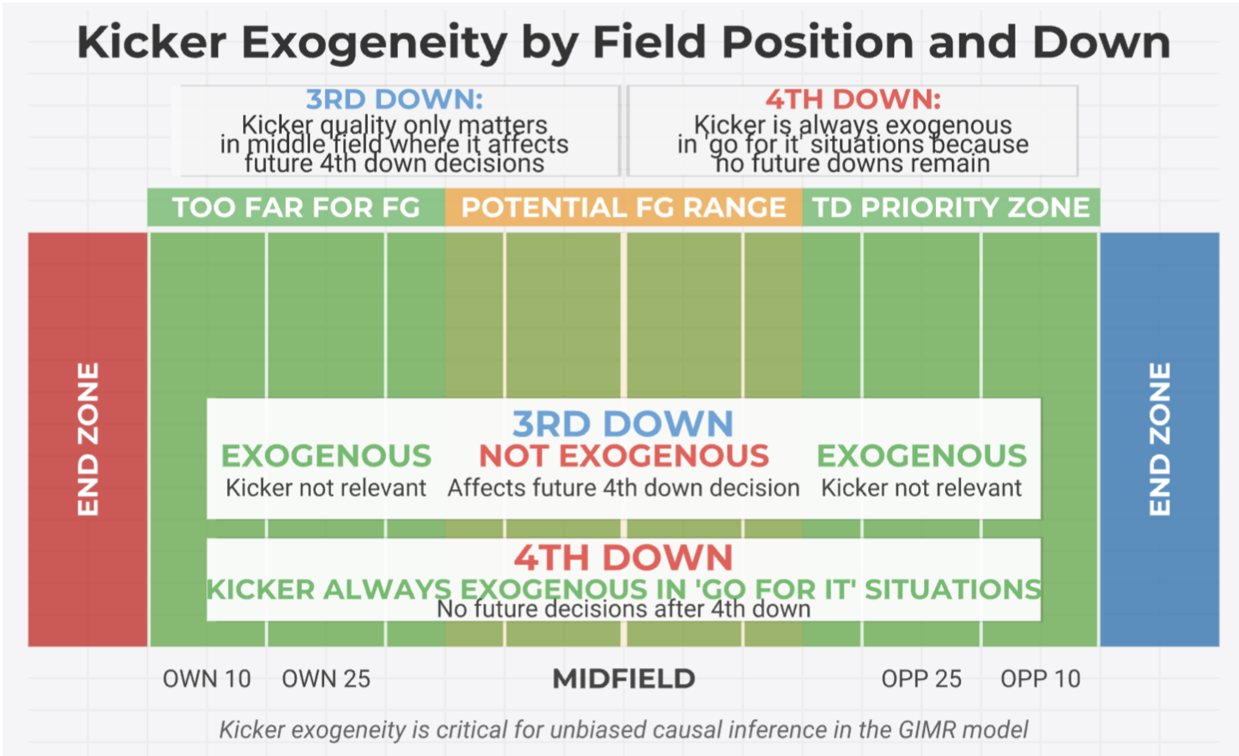


Figure 1: Kicker Plot

attempt probability, they are included. The logic for them remains the same as the kickers and therefore there is no harm in including them in the selection model.

5.2 Generalized Inverse Mills Ratio (GIMR)

In examining the following tables one can see significance of our GIMR is confirmed. This tells us that coaches in the NFL see factors that we cannot and act on them in the proper way. When a coach is more likely to attempt a fourth down conversion attempt that they are then also more likely to convert that fourth down attempt due to unseen variables. The GIMR also allows for bias free interpretation of the marginal effects of variables across models.

5.3 Player Analysis

5.3.1 Pretense

The exogeneity of kicker grades allows a bias-free analysis of individual players and coaches. Players are evaluated in terms of their performance in the 12 week timeframe. This time frame is arbitrary and can be a source of future research with how to optimize the measurement of player values. Players that are used are either "On-Field" or "Starter" players. On-field players are players that are on the field during the fourth down attempt. Starter players are the players that are listed as the starter on the depth chart. In both situations players are sorted via depth chart postings. For each of these categories we then create models that focus on either offensive of defensive players. If an offensive model is ran we select multiple features about each player on the offense while keeping one key feature of each defensive player. The opposite is done for the defensive model.

The offensive and defensive models are both measured in three different ways. For offensive models we measure by PFF grades, completions/receptions and yards. For defensive models we measure by PFF grades, stops and tackles. For our current work we will use PFF grades for our main analysis due to its ability to capture the overall performance of a player in a specific area. The other measures will be kept as robustness checks. To evaluate coaches' decision making, the attempt models are able to show if a coach is properly deciding to attempt a fourth down based on features of players that influence the conversion of a fourth down.

5.3.2 Offensive Player Grades Model Results

The following tables displays the results of the Offensive Player Grades Models across LPM, probit and logit models.

Table 3: Offense Grades Table

	Starter Models			On-Field Models			Attempt Models		
Variable	**Starter LPM**	**Starter Probit**	**Starter Logit**	**On-Field LPM**	**On-Field Probit**	**On-Field Logit**	**Attempt LPM**	**Attempt Probit**	**Attempt Logit**
Quarterbacks									
Short Grades Pass	0.027(0.016)*	0.034(0.019)*	0.035(0.020)*	0.033(0.014)**	0.040(0.016)**	0.042(0.017)**	0.007(0.004)	0.007(0.005)	0.006(0.004)
Medium Grades Pass	-0.028(0.014)**	-0.034(0.016)**	-0.037(0.016)**	-0.013(0.015)	-0.014(0.017)	-0.016(0.017)	0.007(0.004)*	0.005(0.004)	0.004(0.004)
Deep Grades Pass	-0.005(0.013)	-0.006(0.015)	-0.007(0.016)	-0.012(0.013)	-0.015(0.015)	-0.016(0.016)	0.006(0.004)	0.005(0.004)	0.005(0.003)
Running Backs									
Grades Pass Block	-0.016(0.010)	-0.018(0.012)	-0.020(0.012)	0.008(0.010)	0.008(0.011)	0.009(0.012)	-0.004(0.003)	-0.006(0.003)*	-0.005(0.003)*
Grades Run	0.003(0.010)	0.002(0.012)	0.002(0.012)	0.003(0.027)	0.002(0.031)	0.005(0.032)	0.021(0.011)*	0.020(0.012)*	0.018(0.010)*
Short Grades Pass	0.005(0.010)	0.006(0.011)	0.007(0.012)	-0.003(0.010)	-0.003(0.012)	-0.004(0.012)	0.000(0.003)	0.001(0.003)	0.000(0.003)
Medium Grades Pass	-0.013(0.009)	-0.014(0.011)	-0.015(0.011)	-0.006(0.009)	-0.006(0.010)	-0.007(0.010)	-0.001(0.003)	-0.001(0.003)	-0.001(0.003)
Deep Grades Pass	0.003(0.009)	0.003(0.011)	0.003(0.011)	-0.001(0.009)	-0.001(0.010)	-0.001(0.010)	0.003(0.003)	0.002(0.003)	0.002(0.002)
WR1									
Grades Run Block	0.008(0.015)	0.006(0.017)	0.007(0.018)	-0.008(0.021)	-0.010(0.024)	-0.009(0.025)	-0.002(0.004)	-0.002(0.005)	-0.002(0.004)
Short Grades Pass	-0.001(0.013)	-0.003(0.015)	-0.003(0.016)	-0.001(0.017)	-0.004(0.019)	-0.003(0.020)	0.002(0.004)	0.002(0.004)	0.001(0.004)
Medium Grades Pass	0.010(0.013)	0.014(0.015)	0.015(0.015)	0.008(0.014)	0.010(0.016)	0.010(0.017)	0.008(0.004)**	0.007(0.004)	0.007(0.004)*
Deep Grades Pass	-0.003(0.011)	-0.003(0.012)	-0.003(0.012)	0.003(0.012)	0.005(0.014)	0.005(0.014)	0.002(0.003)	0.003(0.003)	0.003(0.003)
WR2									
Grades Run Block	-0.010(0.016)	-0.013(0.019)	-0.013(0.019)	0.016(0.038)	0.020(0.043)	0.019(0.045)	0.004(0.005)	0.004(0.005)	0.005(0.005)
Short Grades Pass	-0.012(0.012)	-0.015(0.014)	-0.016(0.015)	0.005(0.026)	0.007(0.029)	0.006(0.031)	0.009(0.004)**	0.008(0.004)**	0.007(0.004)**
Medium Grades Pass	0.011(0.011)	0.012(0.013)	0.013(0.013)	-0.002(0.020)	-0.006(0.023)	-0.005(0.024)	0.002(0.004)	0.001(0.004)	0.001(0.003)
Deep Grades Pass	0.004(0.010)	0.005(0.011)	0.005(0.012)	0.013(0.016)	0.019(0.018)	0.018(0.019)	0.000(0.003)	0.001(0.003)	0.000(0.003)
WR3									
Grades Run Block	-0.013(0.013)	-0.015(0.016)	-0.017(0.016)	0.024(0.029)	0.026(0.034)	0.027(0.035)	0.001(0.004)	0.002(0.004)	0.002(0.004)
Short Grades Pass	-0.016(0.011)	-0.019(0.012)	-0.021(0.013)	-0.008(0.023)	-0.008(0.026)	-0.007(0.027)	-0.003(0.003)	-0.003(0.004)	-0.002(0.003)
Medium Grades Pass	0.002(0.010)	0.002(0.012)	0.003(0.012)	0.011(0.019)	0.009(0.021)	0.010(0.022)	0.006(0.003)**	0.006(0.003)*	0.005(0.003)*
Deep Grades Pass	0.010(0.010)	0.012(0.011)	0.012(0.011)	0.019(0.015)	0.023(0.017)	0.022(0.018)	0.001(0.003)	0.002(0.003)	0.001(0.003)
TE1									
Grades Pass Block	0.028(0.011)**	0.032(0.013)**	0.034(0.014)**	0.024(0.012)**	0.030(0.013)**	0.031(0.014)**	-0.002(0.004)	-0.002(0.004)	-0.002(0.003)
Grades Run Block	-0.017(0.015)	-0.019(0.017)	-0.020(0.018)	-0.012(0.016)	-0.017(0.018)	-0.017(0.019)	0.002(0.005)	0.003(0.005)	0.001(0.004)
Short Grades Pass	-0.003(0.012)	-0.003(0.014)	-0.005(0.014)	-0.001(0.016)	-0.001(0.018)	-0.000(0.018)	0.001(0.004)	-0.000(0.005)	-0.000(0.004)
Medium Grades Pass	-0.020(0.010)*	-0.024(0.012)**	-0.024(0.012)**	0.020(0.011)*	0.023(0.012)*	0.024(0.012)*	0.004(0.003)	0.003(0.003)	0.002(0.003)
Deep Grades Pass	-0.000(0.010)	-0.001(0.011)	-0.001(0.011)	-0.014(0.009)	-0.016(0.011)	-0.017(0.011)	0.002(0.003)	0.002(0.003)	0.002(0.003)
OL1									
Grades Pass Block	0.015(0.012)	0.018(0.013)	0.019(0.014)	-0.007(0.010)	-0.009(0.011)	-0.008(0.011)	-0.003(0.003)	-0.003(0.004)	-0.003(0.003)
Grades Run Block	0.012(0.014)	0.015(0.015)	0.015(0.016)	0.006(0.010)	0.009(0.012)	0.009(0.012)	0.002(0.004)	0.000(0.004)	0.001(0.004)
OL2									
Grades Pass Block	-0.014(0.012)	-0.017(0.013)	-0.018(0.013)	-0.011(0.010)	-0.012(0.011)	-0.012(0.012)	0.001(0.003)	-0.001(0.003)	-0.001(0.003)
Grades Run Block	0.005(0.014)	0.004(0.016)	0.005(0.017)	-0.009(0.010)	-0.011(0.011)	-0.012(0.012)	0.004(0.004)	0.005(0.004)	0.005(0.004)
OL3									
Grades Pass Block	0.008(0.011)	0.009(0.013)	0.010(0.013)	0.003(0.009)	0.004(0.011)	0.004(0.011)	0.004(0.003)	0.004(0.004)	0.003(0.003)
Grades Run Block	-0.018(0.014)	-0.021(0.016)	-0.022(0.016)	0.010(0.009)	0.014(0.011)	0.014(0.012)	-0.004(0.004)	-0.005(0.004)	-0.005(0.004)
OL4									
Grades Pass Block	0.020(0.012)*	0.024(0.013)*	0.024(0.014)*	0.006(0.010)	0.008(0.011)	0.008(0.011)	-0.001(0.003)	-0.002(0.003)	-0.001(0.003)
Grades Run Block	-0.007(0.013)	-0.008(0.015)	-0.008(0.015)	-0.015(0.009)	-0.018(0.011)*	-0.019(0.011)*	-0.002(0.004)	-0.001(0.004)	-0.001(0.004)
OL5									
Grades Pass Block	0.034(0.012)***	0.039(0.013)***	0.041(0.014)***	-0.007(0.010)	-0.008(0.011)	-0.008(0.012)	0.003(0.003)	0.005(0.004)	0.003(0.003)
Grades Run Block	-0.033(0.013)**	-0.040(0.014)***	-0.040(0.015)***	0.010(0.010)	0.011(0.011)	0.011(0.012)	-0.002(0.004)	-0.003(0.004)	-0.003(0.004)
Perf. Measures									
GIMR	0.040(0.009)***	0.050(0.012)***	0.051(0.013)***	0.037(0.009)***	0.045(0.012)***	0.046(0.013)***			
Control Variables									
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Player Presence	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Control Def. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Off. Players	No	No	No	No	No	No	No	No	No

Note:
Marginal effects reported with standard errors in parentheses and significance levels.
[*] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

It is found that the effect of an on-field quarterback is positively significant across all model types. This finding exemplifies how quarterbacks that excel in passing the ball on quicker routes perform better on fourth downs. Contrary to this a starting quarterback's grades are not significant in the decision making of coaches whether to attempt a fourth down. This is to say that the coaches fail to properly factor in the specific archetypes of quarterback that their team possesses. While the strength of significance with the starter versions of the models is weaker, the following table 4 cannot find a statistically significant difference between the marginal effects of the on-field and starter short passing grades for quarterbacks. However, there is a statistical different between the on-field marginal effects for quarterback grades and the attempt models' marginal effects of quarterback grades.

Similarly we see the Tight End 1 slot shows significance at a p-value of 0.05 in both on-field and starter models in terms of their pass blocking grades. The further lack of any positive significance in terms of a Tight End's pass catching grades tells us that a tight end's ability to block is more important then their ability to catch the ball in fourth down situations. Similarly to Quarterbacks, the coaches fail to properly factor in the specific archetype of tight end that their team possesses as there is a lack of significance within attempt models.

The positive signs of significance of running back grades in our attempt models suggests that coaches will factor in a running back in a fourth down situation. This checks out intuitivly as the ball is often run on fourth down and the running back would appear to be a key player in running the ball. However there is a lack of significance of the running back on the actual conversion of the attempt on fourth down. To support this the offensive linemen we will see show strong joint significance on conversion and not on attempt decisions. This suggests that coaches are overly focusing on the skill of their running back rather then the skill of their linemen.

Table 4: Offense Grades Differences in Significance Table

	LPM			Probit			Logit		
Variable	**Start-Fld**	**Start-Att**	**Fld-Att**	**Start-Fld**	**Start-Att**	**Fld-Att**	**Start-Fld**	**Start-Att**	**Fld-Att**
Quarterbacks									
Short Grades Pass	-0.006(0.021)	0.019(0.017)	0.026(0.014)*	-0.006(0.025)	0.027(0.019)	0.033(0.017)**	-0.007(0.026)	0.029(0.020)	0.036(0.017)**
Medium Grades Pass	-0.016(0.020)	-0.035(0.014)**	-0.019(0.015)	-0.020(0.023)	-0.039(0.016)**	-0.019(0.017)	-0.021(0.024)	-0.041(0.017)**	-0.020(0.018)
Deep Grades Pass	0.007(0.019)	-0.011(0.014)	-0.018(0.014)	0.009(0.022)	-0.011(0.016)	-0.020(0.016)	0.009(0.022)	-0.012(0.016)	-0.021(0.016)
Running Backs									
Grades Pass Block	-0.024(0.014)*	-0.011(0.011)	0.012(0.010)	-0.027(0.016)	-0.013(0.012)	0.014(0.012)	-0.028(0.017)*	-0.015(0.012)	0.014(0.012)
Grades Run	0.000(0.029)	-0.019(0.015)	-0.019(0.029)	0.000(0.033)	-0.018(0.017)	-0.018(0.033)	-0.002(0.034)	-0.015(0.016)	-0.013(0.033)
Short Grades Pass	0.008(0.014)	0.005(0.010)	-0.003(0.011)	0.009(0.016)	0.006(0.012)	-0.003(0.012)	0.010(0.017)	0.006(0.012)	-0.004(0.012)
Medium Grades Pass	-0.007(0.013)	-0.012(0.010)	-0.005(0.009)	-0.008(0.015)	-0.013(0.011)	-0.005(0.010)	-0.008(0.015)	-0.014(0.011)	-0.006(0.011)
Deep Grades Pass	0.003(0.013)	-0.000(0.010)	-0.004(0.009)	0.004(0.015)	0.001(0.011)	-0.003(0.010)	0.004(0.015)	0.001(0.011)	-0.002(0.011)
WR1									
Grades Run Block	0.015(0.026)	0.010(0.016)	-0.005(0.021)	0.015(0.029)	0.008(0.018)	-0.007(0.024)	0.015(0.031)	0.009(0.018)	-0.006(0.025)
Short Grades Pass	0.000(0.021)	-0.003(0.014)	-0.003(0.017)	0.001(0.024)	-0.005(0.016)	-0.005(0.019)	0.000(0.025)	-0.004(0.016)	-0.005(0.020)
Medium Grades Pass	0.002(0.019)	0.002(0.014)	-0.000(0.015)	0.004(0.022)	0.007(0.016)	0.003(0.017)	0.005(0.023)	0.008(0.016)	0.003(0.017)
Deep Grades Pass	-0.006(0.016)	-0.005(0.011)	0.001(0.013)	-0.008(0.018)	-0.006(0.012)	0.002(0.014)	-0.008(0.019)	-0.006(0.013)	0.002(0.014)
WR2									
Grades Run Block	-0.026(0.041)	-0.015(0.017)	0.012(0.038)	-0.032(0.047)	-0.017(0.019)	0.015(0.044)	-0.032(0.049)	-0.018(0.020)	0.014(0.045)
Short Grades Pass	-0.017(0.028)	-0.021(0.013)*	-0.004(0.026)	-0.022(0.033)	-0.023(0.015)	-0.001(0.030)	-0.022(0.034)	-0.023(0.015)	-0.001(0.031)
Medium Grades Pass	0.013(0.023)	0.009(0.012)	-0.004(0.020)	0.019(0.026)	0.011(0.013)	-0.007(0.023)	0.018(0.027)	0.012(0.013)	-0.006(0.024)
Deep Grades Pass	-0.008(0.019)	0.004(0.011)	0.012(0.016)	-0.013(0.021)	0.005(0.012)	0.018(0.018)	-0.013(0.022)	0.005(0.012)	0.018(0.019)
WR3									
Grades Run Block	-0.036(0.032)	-0.014(0.014)	0.022(0.029)	-0.041(0.037)	-0.017(0.016)	0.024(0.034)	-0.044(0.038)	-0.019(0.016)	0.025(0.035)
Short Grades Pass	-0.008(0.025)	-0.013(0.011)	-0.005(0.023)	-0.011(0.029)	-0.016(0.013)	-0.005(0.026)	-0.014(0.030)	-0.018(0.013)	-0.004(0.027)
Medium Grades Pass	-0.009(0.021)	-0.004(0.011)	0.005(0.019)	-0.007(0.024)	-0.003(0.012)	0.004(0.021)	-0.007(0.025)	-0.002(0.012)	0.005(0.022)
Deep Grades Pass	-0.009(0.018)	0.009(0.010)	0.018(0.015)	-0.011(0.020)	0.010(0.011)	0.021(0.017)	-0.010(0.021)	0.011(0.012)	0.021(0.018)
TE1									
Grades Pass Block	0.005(0.016)	0.030(0.012)**	0.025(0.012)**	0.002(0.019)	0.034(0.014)**	0.032(0.014)**	0.003(0.019)	0.036(0.014)**	0.032(0.014)**
Grades Run Block	-0.006(0.022)	-0.019(0.016)	-0.013(0.016)	-0.002(0.025)	-0.021(0.018)	-0.019(0.019)	-0.003(0.026)	-0.021(0.018)	-0.018(0.019)
Short Grades Pass	-0.002(0.020)	-0.003(0.013)	-0.001(0.016)	-0.003(0.022)	-0.003(0.015)	-0.000(0.018)	-0.005(0.023)	-0.005(0.015)	0.000(0.019)
Medium Grades Pass	-0.040(0.015)***	-0.024(0.011)**	0.016(0.011)	-0.046(0.017)***	-0.027(0.012)**	0.019(0.012)	-0.048(0.017)***	-0.026(0.012)**	0.022(0.013)*
Deep Grades Pass	0.014(0.013)	-0.002(0.010)	-0.016(0.010)	0.015(0.015)	-0.004(0.011)	-0.018(0.011)*	0.016(0.016)	-0.003(0.012)	-0.019(0.011)*
OL1									
Grades Pass Block	0.022(0.015)	0.018(0.012)	-0.004(0.010)	0.027(0.017)	0.021(0.014)	-0.006(0.011)	0.027(0.018)	0.021(0.014)	-0.006(0.012)
Grades Run Block	0.005(0.017)	0.010(0.014)	0.005(0.011)	0.007(0.019)	0.015(0.016)	0.008(0.012)	0.006(0.020)	0.015(0.016)	0.009(0.013)
OL2									
Grades Pass Block	-0.003(0.015)	-0.015(0.012)	-0.011(0.010)	-0.005(0.017)	-0.016(0.013)	-0.011(0.012)	-0.006(0.018)	-0.017(0.014)	-0.011(0.012)
Grades Run Block	0.014(0.017)	0.002(0.015)	-0.013(0.010)	0.015(0.020)	-0.000(0.017)	-0.016(0.012)	0.018(0.021)	0.001(0.017)	-0.017(0.012)
OL3									
Grades Pass Block	0.005(0.015)	0.004(0.012)	-0.001(0.010)	0.005(0.017)	0.005(0.013)	-0.000(0.011)	0.007(0.017)	0.007(0.013)	0.001(0.012)
Grades Run Block	-0.028(0.017)*	-0.014(0.015)	0.014(0.010)	-0.035(0.019)*	-0.016(0.016)	0.019(0.012)	-0.036(0.020)*	-0.017(0.017)	0.019(0.012)
OL4									
Grades Pass Block	0.014(0.015)	0.021(0.012)*	0.007(0.010)	0.016(0.017)	0.026(0.014)*	0.009(0.011)	0.017(0.018)	0.026(0.014)*	0.009(0.012)
Grades Run Block	0.008(0.016)	-0.005(0.014)	-0.013(0.010)	0.010(0.018)	-0.007(0.015)	-0.017(0.011)	0.011(0.019)	-0.007(0.016)	-0.018(0.011)
OL5									
Grades Pass Block	0.041(0.015)***	0.031(0.012)**	-0.010(0.011)	0.047(0.017)***	0.034(0.014)**	-0.013(0.012)	0.048(0.018)***	0.037(0.014)***	-0.011(0.012)
Grades Run Block	-0.043(0.016)***	-0.031(0.014)**	0.012(0.011)	-0.051(0.018)***	-0.037(0.015)**	0.013(0.012)	-0.051(0.019)***	-0.038(0.015)**	0.013(0.012)

Note:
Differences shown with standard errors in parentheses. Positive values indicate the first definition has a larger effect than the second.
* * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The following tables display joint significance of the Offensive Player Grades Models across LPM, probit and logit models. Here they are grouped by position groups. This analysis is useful in the case of where a position has multiple players that are on the field of that position. For example, the offensive line has 5 players that are on the field at once. In our data we generalize the positions of these players from positions like left tackle, left guard, center, right guard and right tackle to avoid data loss when matching players into there designated slots. For example if an on-field left tackle where to play a snap at left guard this would create a missing value when we look to match a player to their respective column. This same logic applies for the defensive players as well.

While this generalization creates a loss of position tracking in our individual models, we still can use the joint significance to extract meaningful causal inferences. For example we can observe how the Starter models find signifiance with the offensive line as a whole. The offensive line grade's are missed by coaches in their attempt models.

Table 5: Offense Grades Model Results

	Starter Models			On-Field Models			Attempt Models		
Variable	**Starter LPM**	**Starter Probit**	**Starter Logit**	**On-Field LPM**	**On-Field Probit**	**On-Field Logit**	**Attempt LPM**	**Attempt Probit**	**Attempt Logit**
Quarterbacks	2.03* [4,4015]	2.48** [4,4015]	2.78** [4,4015]	2.44* [3,4016]	2.67** [3,4016]	2.81** [3,4016]	2.12* [4,25604]	1.39 [4,25604]	1.47 [4,25604]
Running Backs	0.82 [6,4013]	0.82 [6,4013]	0.87 [6,4013]	0.21 [6,4013]	0.17 [6,4013]	0.21 [6,4013]	1.83* [6,25602]	1.51 [6,25602]	1.47 [6,25602]
Wide Receivers	0.68 [15,4004]	0.73 [15,4004]	0.77 [15,4004]	0.29 [15,4004]	0.34 [15,4004]	0.31 [15,4004]	1.23 [15,25593]	0.92 [15,25593]	1.05 [15,25593]
Tight Ends	2.74** [6,4013]	2.83*** [6,4013]	2.87*** [6,4013]	1.80* [6,4013]	2.00* [6,4013]	2.00* [6,4013]	0.37 [6,25602]	0.39 [6,25602]	0.31 [6,25602]
Offensive Line	2.45*** [10,4009]	2.71*** [10,4009]	2.67*** [10,4009]	0.86 [10,4009]	0.98 [10,4009]	0.96 [10,4009]	0.53 [10,25598]	0.73 [10,25598]	0.67 [10,25598]
Control Variables									
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GIMR	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Player Presence	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Control Def. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Off. Players	No	No	No	No	No	No	No	No	No

Note:
F-statistics reported with degrees of freedom [df1,df2] and significance levels.
* * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In our last set of offensive F-values tables, we find that our starter models show significance for our entire offense as a whole. This is consistent across our robustness checks when models are ran with yards and completions as well. Attempt models then show no indication of the offensive players jointly having a influence on the attempt of a fourth down.

Table 6: Joint Significance Offense

	Starter Models			On-Field Models			Attempt Models		
Dataset	**Starter LPM**	**Starter Probit**	**Starter Logit**	**On-Field LPM**	**On-Field Probit**	**On-Field Logit**	**Attempt LPM**	**Attempt Probit**	**Attempt Logit**
Offense Grades	1.56** [41,3978]	1.70*** [41,3978]	1.75*** [41,3978]	0.81 [40,3979]	0.90 [40,3979]	0.90 [40,3979]	1.11 [41,25567]	0.93 [41,25567]	0.95 [41,25567]
Offense Yards	2.00*** [41,3978]	2.19*** [41,3978]	2.17*** [41,3978]	0.85 [40,3979]	0.92 [40,3979]	0.91 [40,3979]	0.79 [41,25567]	0.91 [41,25567]	0.84 [41,25567]
Offense Completions	1.86*** [41,3978]	2.02*** [41,3978]	2.03*** [41,3978]	1.01 [40,3979]	1.11 [40,3979]	1.08 [40,3979]	0.65 [41,25567]	0.72 [41,25567]	0.65 [41,25567]
Control Variables									
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GIMR	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Player Presence	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Control Def. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Off. Players	No	No	No	No	No	No	No	No	No

Note:
F-statistics reported with degrees of freedom [df1,df2] and significance levels for ALL position variables combined.
* * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3.3 Defensive Player Grades Model Results

Our defensive players individually suffer relatively more from positions being generalized. This is caused by how all positions on defense have many but small differences. However one result can be found with the starting linebackers of a team in a model that measures the effect of linebackers zone coverage grades on the decision to attempt a fourth down. The LB1 and LB2 both have nearly identical negative marginal effects in this matter. With coaches allowing this effect on their decision making it is then concerning to find that the LB1 and LB2 zone coverage grades are not significant in the actual conversion of a fourth down attempt.

Table 7: Defense Grades Table

	Starter Models			On-Field Models			Attempt Models		
Variable	**Starter LPM**	**Starter Probit**	**Starter Logit**	**On-Field LPM**	**On-Field Probit**	**On-Field Logit**	**Attempt LPM**	**Attempt Probit**	**Attempt Logit**
DL1									
Grades Run Defense	0.034(0.012)***	0.039(0.014)***	0.040(0.014)***	-0.012(0.010)	-0.015(0.012)	-0.015(0.012)	-0.002(0.004)	-0.004(0.004)	-0.003(0.004)
Grds Pass Rush Def	-0.011(0.012)	-0.011(0.013)	-0.011(0.014)	-0.009(0.010)	-0.009(0.012)	-0.010(0.012)	-0.006(0.004)	-0.004(0.004)	-0.004(0.004)
DL2									
Grades Run Defense	-0.004(0.013)	-0.005(0.014)	-0.004(0.015)	0.014(0.020)	0.018(0.023)	0.019(0.023)	0.004(0.004)	0.006(0.004)	0.005(0.004)
Grds Pass Rush Def	0.013(0.012)	0.014(0.013)	0.014(0.014)	-0.013(0.020)	-0.017(0.023)	-0.017(0.023)	-0.001(0.004)	-0.000(0.004)	-0.000(0.003)
DL3									
Grades Run Defense	-0.021(0.013)*	-0.026(0.014)*	-0.026(0.015)*	0.006(0.037)	0.005(0.041)	0.008(0.042)	-0.000(0.004)	0.001(0.004)	-0.000(0.004)
Grds Pass Rush Def	0.007(0.012)	0.010(0.014)	0.010(0.014)	0.006(0.037)	0.009(0.041)	0.006(0.042)	-0.001(0.004)	-0.001(0.004)	-0.001(0.004)
DL4									
Grades Run Defense	-0.009(0.013)	-0.009(0.015)	-0.010(0.015)	-0.014(0.046)	-0.020(0.051)	-0.019(0.053)	0.003(0.004)	0.001(0.004)	0.001(0.004)
Grds Pass Rush Def	-0.002(0.013)	-0.002(0.015)	-0.002(0.015)	-0.015(0.047)	-0.015(0.052)	-0.016(0.053)	-0.001(0.004)	0.002(0.004)	0.001(0.004)
LB1									
Grades Run Defense	-0.011(0.011)	-0.014(0.013)	-0.015(0.014)	0.008(0.010)	0.009(0.012)	0.010(0.012)	0.002(0.004)	0.002(0.004)	0.001(0.004)
Man Grades Cov Def	-0.003(0.009)	-0.004(0.010)	-0.004(0.011)	-0.001(0.008)	-0.000(0.010)	-0.001(0.010)	0.004(0.003)	0.003(0.003)	0.003(0.003)
Zone Grades Cov Def	-0.004(0.011)	-0.005(0.012)	-0.005(0.013)	-0.005(0.010)	-0.007(0.011)	-0.008(0.012)	-0.007(0.003)**	-0.007(0.003)**	-0.006(0.003)*
LB2									
Grades Run Defense	0.033(0.011)***	0.040(0.013)***	0.041(0.013)***	-0.010(0.022)	-0.013(0.025)	-0.015(0.026)	0.004(0.004)	0.005(0.004)	0.005(0.003)
Man Grades Cov Def	-0.010(0.009)	-0.012(0.010)	-0.013(0.011)	-0.005(0.013)	-0.005(0.015)	-0.006(0.016)	-0.002(0.003)	-0.002(0.003)	-0.002(0.002)
Zone Grades Cov Def	0.001(0.010)	0.001(0.011)	0.001(0.012)	0.020(0.018)	0.024(0.020)	0.025(0.021)	-0.008(0.003)**	-0.007(0.003)**	-0.006(0.003)**
LB3									
Grades Run Defense	-0.001(0.011)	-0.000(0.013)	-0.001(0.013)	-0.069(0.045)	-0.084(0.052)	-0.080(0.054)	-0.004(0.004)	-0.003(0.004)	-0.003(0.004)
Man Grades Cov Def	-0.013(0.009)	-0.015(0.011)	-0.015(0.011)	0.020(0.018)	0.022(0.020)	0.024(0.021)	0.001(0.003)	0.002(0.003)	0.002(0.003)
Zone Grades Cov Def	0.014(0.011)	0.016(0.012)	0.017(0.013)	-0.049(0.028)*	-0.058(0.033)*	-0.062(0.035)*	-0.000(0.004)	-0.001(0.004)	-0.000(0.003)
LB4									
Grades Run Defense	0.012(0.011)	0.014(0.013)	0.014(0.014)	0.036(0.054)	0.044(0.061)	0.042(0.064)	-0.006(0.005)	-0.007(0.005)	-0.007(0.004)
Man Grades Cov Def	0.010(0.011)	0.013(0.013)	0.013(0.013)	0.005(0.018)	0.005(0.020)	0.006(0.021)	0.002(0.003)	0.001(0.003)	0.002(0.003)
Zone Grades Cov Def	-0.004(0.012)	-0.007(0.014)	-0.007(0.014)	-0.008(0.028)	-0.004(0.034)	-0.003(0.036)	0.005(0.004)	0.008(0.004)**	0.006(0.003)**
CB1									
Grades Run Defense	0.037(0.013)***	0.043(0.015)***	0.045(0.015)***	0.004(0.018)	0.005(0.019)	0.005(0.020)	-0.001(0.005)	-0.002(0.005)	-0.001(0.004)
Man Grades Cov Def	-0.011(0.011)	-0.014(0.013)	-0.014(0.013)	-0.037(0.013)***	-0.041(0.015)***	-0.043(0.016)***	0.002(0.004)	0.002(0.004)	0.002(0.003)
Zone Grades Cov Def	-0.015(0.012)	-0.016(0.014)	-0.017(0.014)	-0.028(0.016)*	-0.034(0.017)*	-0.034(0.018)*	-0.000(0.004)	0.000(0.004)	-0.000(0.004)
CB2									
Grades Run Defense	-0.044(0.012)***	-0.051(0.014)***	-0.054(0.015)***	-0.025(0.021)	-0.030(0.023)	-0.031(0.024)	-0.002(0.004)	-0.002(0.005)	-0.002(0.004)
Man Grades Cov Def	-0.019(0.011)*	-0.020(0.013)	-0.022(0.013)*	0.001(0.018)	0.002(0.021)	0.002(0.022)	-0.001(0.004)	-0.002(0.004)	-0.001(0.003)
Zone Grades Cov Def	0.023(0.012)*	0.025(0.014)*	0.026(0.014)*	-0.000(0.022)	0.001(0.025)	0.001(0.026)	0.001(0.004)	0.000(0.004)	-0.000(0.003)
CB3									
Grades Run Defense	0.006(0.011)	0.007(0.013)	0.008(0.013)	-0.025(0.036)	-0.030(0.041)	-0.030(0.042)	-0.006(0.004)	-0.004(0.004)	-0.004(0.004)
Man Grades Cov Def	0.006(0.011)	0.005(0.013)	0.006(0.013)	-0.025(0.029)	-0.029(0.033)	-0.027(0.034)	0.005(0.003)	0.006(0.004)	0.005(0.003)
Zone Grades Cov Def	-0.007(0.012)	-0.008(0.014)	-0.009(0.014)	0.008(0.035)	0.009(0.038)	0.008(0.040)	-0.004(0.004)	-0.006(0.004)	-0.005(0.004)
S1									
Grades Run Defense	-0.027(0.014)**	-0.030(0.015)*	-0.032(0.016)**	-0.009(0.015)	-0.010(0.016)	-0.011(0.017)	-0.003(0.005)	-0.005(0.006)	-0.002(0.005)
Man Grades Cov Def	0.007(0.013)	0.008(0.014)	0.008(0.015)	0.004(0.012)	0.004(0.014)	0.004(0.015)	-0.002(0.005)	-0.002(0.005)	-0.001(0.004)
Zone Grades Cov Def	0.013(0.012)	0.014(0.014)	0.014(0.014)	0.013(0.012)	0.016(0.014)	0.017(0.015)	0.001(0.004)	-0.001(0.005)	-0.000(0.004)
S2									
Grades Run Defense	-0.026(0.013)**	-0.032(0.015)**	-0.032(0.016)**	-0.028(0.029)	-0.031(0.034)	-0.032(0.036)	0.006(0.006)	0.005(0.006)	0.004(0.005)
Man Grades Cov Def	0.015(0.013)	0.018(0.015)	0.018(0.015)	-0.001(0.023)	-0.002(0.027)	-0.002(0.028)	0.006(0.004)	0.008(0.005)*	0.007(0.004)*
Zone Grades Cov Def	-0.001(0.014)	-0.000(0.015)	-0.001(0.016)	-0.018(0.025)	-0.020(0.029)	-0.021(0.030)	0.003(0.004)	0.006(0.005)	0.006(0.004)
S3									
Grades Run Defense	0.011(0.012)	0.014(0.014)	0.013(0.014)	0.013(0.041)	0.016(0.047)	0.015(0.049)	0.001(0.004)	0.003(0.004)	0.003(0.003)
Man Grades Cov Def	-0.020(0.012)*	-0.024(0.013)*	-0.024(0.014)*	-0.011(0.042)	-0.006(0.049)	-0.009(0.052)	-0.005(0.004)	-0.005(0.004)	-0.005(0.003)
Zone Grades Cov Def	0.001(0.013)	0.000(0.015)	0.000(0.015)	0.004(0.043)	-0.005(0.051)	-0.001(0.055)	0.002(0.004)	0.001(0.004)	0.002(0.004)
Perf. Measures									
GIMR	0.048(0.010)***	0.055(0.013)***	0.057(0.013)***	0.043(0.010)***	0.049(0.013)***	0.051(0.014)***			
Control Variables									
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Player Presence	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Control Off. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Def. Players	No	No	No	No	No	No	No	No	No

Note:

Marginal effects reported with standard errors in parentheses and significance levels.

[*] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Defense Grades Differences in Significance Table

	LPM			Probit			Logit		
Variable	**Start-Fld**	**Start-Att**	**Fld-Att**	**Start-Fld**	**Start-Att**	**Fld-Att**	**Start-Fld**	**Start-Att**	**Fld-Att**
DL1									
Grades Run Defense	0.046(0.016)***	0.035(0.013)***	-0.011(0.011)	0.054(0.018)***	0.043(0.015)***	-0.012(0.013)	0.055(0.019)***	0.043(0.015)***	-0.013(0.013)
Grds Pass Rush Def	-0.002(0.016)	-0.005(0.012)	-0.003(0.011)	-0.002(0.018)	-0.007(0.014)	-0.005(0.012)	-0.001(0.018)	-0.007(0.014)	-0.006(0.013)
DL2									
Grades Run Defense	-0.018(0.024)	-0.008(0.013)	0.010(0.021)	-0.023(0.027)	-0.011(0.015)	0.012(0.023)	-0.023(0.028)	-0.009(0.015)	0.014(0.024)
Grds Pass Rush Def	0.025(0.023)	0.013(0.013)	-0.012(0.020)	0.031(0.026)	0.014(0.014)	-0.017(0.023)	0.031(0.027)	0.014(0.014)	-0.017(0.024)
DL3									
Grades Run Defense	-0.027(0.039)	-0.021(0.013)	0.006(0.037)	-0.031(0.043)	-0.027(0.015)*	0.003(0.041)	-0.033(0.045)	-0.026(0.015)*	0.008(0.042)
Grds Pass Rush Def	0.001(0.039)	0.008(0.013)	0.007(0.037)	0.001(0.043)	0.011(0.014)	0.010(0.041)	0.003(0.044)	0.011(0.014)	0.007(0.042)
DL4									
Grades Run Defense	0.005(0.048)	-0.011(0.013)	-0.017(0.046)	0.010(0.053)	-0.011(0.015)	-0.021(0.051)	0.010(0.055)	-0.011(0.016)	-0.021(0.053)
Grds Pass Rush Def	0.013(0.049)	-0.001(0.014)	-0.014(0.047)	0.013(0.054)	-0.004(0.015)	-0.017(0.052)	0.014(0.055)	-0.003(0.016)	-0.018(0.053)
LB1									
Grades Run Defense	-0.019(0.015)	-0.013(0.012)	0.006(0.011)	-0.023(0.018)	-0.015(0.014)	0.008(0.013)	-0.024(0.018)	-0.016(0.014)	0.008(0.013)
Man Grades Cov Def	-0.001(0.012)	-0.006(0.009)	-0.005(0.009)	-0.004(0.014)	-0.007(0.011)	-0.004(0.010)	-0.004(0.015)	-0.007(0.011)	-0.004(0.011)
Zone Grades Cov Def	0.002(0.015)	0.003(0.011)	0.002(0.011)	0.002(0.017)	0.002(0.013)	-0.000(0.012)	0.003(0.018)	0.001(0.013)	-0.002(0.012)
LB2									
Grades Run Defense	0.044(0.024)*	0.029(0.012)**	-0.014(0.022)	0.053(0.028)*	0.035(0.013)***	-0.018(0.025)	0.056(0.029)*	0.036(0.014)***	-0.020(0.026)
Man Grades Cov Def	-0.005(0.016)	-0.008(0.009)	-0.004(0.013)	-0.007(0.018)	-0.010(0.011)	-0.003(0.016)	-0.007(0.019)	-0.011(0.011)	-0.004(0.016)
Zone Grades Cov Def	-0.019(0.021)	0.009(0.011)	0.028(0.018)	-0.023(0.023)	0.009(0.012)	0.032(0.021)	-0.023(0.024)	0.007(0.012)	0.031(0.021)
LB3									
Grades Run Defense	0.068(0.047)	0.004(0.012)	-0.064(0.045)	0.084(0.053)	0.003(0.013)	-0.081(0.052)	0.079(0.056)	0.002(0.014)	-0.077(0.054)
Man Grades Cov Def	-0.033(0.020)	-0.013(0.010)	0.019(0.018)	-0.037(0.023)	-0.017(0.011)	0.020(0.021)	-0.039(0.024)	-0.018(0.011)	0.021(0.021)
Zone Grades Cov Def	0.063(0.030)**	0.014(0.012)	-0.049(0.028)*	0.074(0.036)**	0.017(0.013)	-0.057(0.034)*	0.079(0.037)**	0.017(0.013)	-0.062(0.035)*
LB4									
Grades Run Defense	-0.024(0.055)	0.019(0.013)	0.043(0.054)	-0.030(0.063)	0.021(0.014)	0.051(0.061)	-0.028(0.065)	0.021(0.014)	0.049(0.064)
Man Grades Cov Def	0.005(0.021)	0.008(0.011)	0.004(0.018)	0.008(0.024)	0.012(0.013)	0.004(0.021)	0.007(0.025)	0.012(0.013)	0.005(0.021)
Zone Grades Cov Def	0.004(0.031)	-0.009(0.013)	-0.013(0.029)	-0.003(0.037)	-0.015(0.015)	-0.012(0.034)	-0.003(0.039)	-0.013(0.015)	-0.010(0.036)
CB1									
Grades Run Defense	0.034(0.022)	0.039(0.013)***	0.005(0.018)	0.038(0.024)	0.045(0.015)***	0.007(0.020)	0.040(0.025)	0.046(0.016)***	0.006(0.020)
Man Grades Cov Def	0.026(0.018)	-0.012(0.012)	-0.039(0.014)***	0.027(0.020)	-0.016(0.014)	-0.043(0.016)***	0.029(0.021)	-0.016(0.014)	-0.045(0.016)***
Zone Grades Cov Def	0.013(0.020)	-0.015(0.013)	-0.028(0.016)*	0.018(0.022)	-0.016(0.015)	-0.034(0.018)*	0.017(0.023)	-0.017(0.015)	-0.033(0.019)*
CB2									
Grades Run Defense	-0.019(0.024)	-0.042(0.013)***	-0.023(0.021)	-0.020(0.028)	-0.048(0.015)***	-0.028(0.024)	-0.022(0.029)	-0.052(0.016)***	-0.029(0.025)
Man Grades Cov Def	-0.020(0.021)	-0.017(0.012)	0.003(0.018)	-0.022(0.024)	-0.019(0.013)	0.003(0.021)	-0.024(0.025)	-0.021(0.014)	0.003(0.022)
Zone Grades Cov Def	0.023(0.025)	0.022(0.012)*	-0.001(0.022)	0.024(0.029)	0.024(0.014)*	0.001(0.026)	0.025(0.030)	0.026(0.015)*	0.001(0.027)
CB3									
Grades Run Defense	0.030(0.038)	0.012(0.012)	-0.018(0.036)	0.037(0.042)	0.011(0.013)	-0.026(0.041)	0.038(0.044)	0.012(0.014)	-0.026(0.042)
Man Grades Cov Def	0.031(0.031)	0.001(0.012)	-0.030(0.029)	0.034(0.035)	-0.000(0.013)	-0.035(0.033)	0.033(0.037)	0.001(0.014)	-0.032(0.034)
Zone Grades Cov Def	-0.015(0.037)	-0.003(0.013)	0.011(0.035)	-0.017(0.041)	-0.002(0.014)	0.015(0.039)	-0.016(0.042)	-0.004(0.015)	0.012(0.040)
S1									
Grades Run Defense	-0.018(0.020)	-0.024(0.015)*	-0.006(0.015)	-0.020(0.023)	-0.025(0.016)	-0.005(0.017)	-0.021(0.023)	-0.030(0.017)*	-0.009(0.018)
Man Grades Cov Def	0.003(0.018)	0.009(0.013)	0.006(0.013)	0.004(0.020)	0.010(0.015)	0.005(0.015)	0.004(0.021)	0.010(0.015)	0.005(0.016)
Zone Grades Cov Def	-0.000(0.018)	0.012(0.013)	0.012(0.013)	-0.002(0.020)	0.015(0.015)	0.017(0.015)	-0.002(0.021)	0.015(0.015)	0.017(0.015)
S2									
Grades Run Defense	0.001(0.032)	-0.032(0.014)**	-0.033(0.030)	-0.001(0.037)	-0.037(0.016)**	-0.036(0.035)	0.001(0.039)	-0.036(0.017)**	-0.036(0.036)
Man Grades Cov Def	0.016(0.027)	0.009(0.014)	-0.008(0.024)	0.020(0.031)	0.010(0.015)	-0.010(0.028)	0.020(0.032)	0.011(0.016)	-0.009(0.029)
Zone Grades Cov Def	0.017(0.028)	-0.004(0.014)	-0.021(0.025)	0.020(0.032)	-0.007(0.016)	-0.027(0.029)	0.020(0.034)	-0.007(0.016)	-0.027(0.030)
S3									
Grades Run Defense	-0.002(0.043)	0.010(0.013)	0.011(0.041)	-0.002(0.049)	0.010(0.014)	0.013(0.047)	-0.003(0.051)	0.010(0.015)	0.013(0.049)
Man Grades Cov Def	-0.009(0.044)	-0.016(0.012)	-0.007(0.043)	-0.018(0.051)	-0.020(0.014)	-0.002(0.049)	-0.015(0.053)	-0.019(0.014)	-0.004(0.052)
Zone Grades Cov Def	-0.003(0.045)	-0.001(0.013)	0.002(0.044)	0.005(0.053)	-0.001(0.015)	-0.006(0.052)	0.002(0.057)	-0.001(0.015)	-0.003(0.055)

Note:
Differences shown with standard errors in parentheses. Positive values indicate the first definition has a larger effect than the second.
[*] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The joint significance, found in the following table, of defensive players by position also shows significance with linebackers in our attempt models. This significance is once again not picked up in any way by our starter or on-field models. However Cornerbacks are significant at a p-value of 0.01 in both on-field and starter models. Finally, There is no evidence of cornerbacks effecting attempt probability.

Table 9: Defense Grades Model Results by Position Groups

	Starter Models			On-Field Models			Attempt Models		
Variable	**Starter LPM**	**Starter Probit**	**Starter Logit**	**On-Field LPM**	**On-Field Probit**	**On-Field Logit**	**Attempt LPM**	**Attempt Probit**	**Attempt Logit**
Defensive Line	1.64 [8,4011]	1.74* [8,4011]	1.66 [8,4011]	0.41 [8,4011]	0.46 [8,4011]	0.47 [8,4011]	0.51 [8,25600]	0.52 [8,25600]	0.52 [8,25600]
Linebackers	1.21 [16,4003]	1.30 [16,4003]	1.32 [16,4003]	0.66 [16,4003]	0.70 [16,4003]	0.68 [16,4003]	1.18 [16,25592]	1.31 [16,25592]	1.32 [16,25592]
Cornerbacks	3.37*** [12,4007]	3.31*** [12,4007]	3.44*** [12,4007]	1.74* [12,4007]	2.13** [12,4007]	2.16** [12,4007]	0.52 [12,25596]	0.57 [12,25596]	0.59 [12,25596]
Safeties	1.39 [11,4008]	1.46 [11,4008]	1.40 [11,4008]	0.53 [11,4008]	0.59 [11,4008]	0.57 [11,4008]	0.65 [11,25597]	0.86 [11,25597]	0.91 [11,25597]
Control Variables									
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GIMR	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Player Presence	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Control Off. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Def. Players	No	No	No	No	No	No	No	No	No

Note:

F-statistics reported with degrees of freedom [df1,df2] and significance levels.

* * p < 0.10, ** p < 0.05, *** p < 0.01

Across our different robustness checks models we find that the defensive players as a whole are significant in our starter models that measure with grades and stops. This is not the case for our on-field models or our tackles model. While our tackles models lacking significance is not surprising, as it is a different type of statistic compared to stops and grades, we find once again that on-field measurements lack significance.

Table 10: Joint Significance Defense

	Starter Models			On-Field Models			Attempt Models		
Dataset	**Starter LPM**	**Starter Probit**	**Starter Logit**	**On-Field LPM**	**On-Field Probit**	**On-Field Logit**	**Attempt LPM**	**Attempt Probit**	**Attempt Logit**
Defense Grades	1.87*** [47,3972]	1.93*** [47,3972]	1.94*** [47,3972]	0.86 [47,3972]	1.00 [47,3972]	1.00 [47,3972]	0.77 [47,25561]	0.88 [47,25561]	0.90 [47,25561]
Defense Stops	1.15 [47,3972]	1.20 [47,3972]	1.22 [47,3972]	0.96 [47,3972]	1.03 [47,3972]	1.07 [47,3972]	0.97 [47,25561]	0.90 [47,25561]	0.92 [47,25561]
Defense Tackles	0.67 [47,3972]	0.71 [47,3972]	0.71 [47,3972]	1.03 [47,3972]	1.13 [47,3972]	1.15 [47,3972]	0.83 [47,25561]	0.79 [47,25561]	0.81 [47,25561]
Control Variables									
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Team Stats	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GIMR	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Player Presence	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Control Off. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Def. Players	No	No	No	No	No	No	No	No	No

Note:

F-statistics reported with degrees of freedom [df1,df2] and significance levels for ALL position variables combined.

* * p < 0.10, ** p < 0.05, *** p < 0.01

The interpretation of starter and on-field models leads us to ask why there is a difference in quality of statistical significance. In our F-Statistic tables starter models appear to capture the performance of positions and teams better. Individual player features also seem to measure player performance better as show by a larger amount of significant variables. This leads us to believe that starter models have less noise in measurement then on-field models due to a higher consistency of players. On-field measurements are subject to players being subbed in and out of the game. While it provides greater detail in measurement it makes player slots have more variance in terms of which players are present where. For example the LB2 of a team can be slotted into either LB slot 1 or 2 depending on the availability of the LB1 in an on-field play. However for a given game that LB 2 player will stay in the second slot for the entire game.

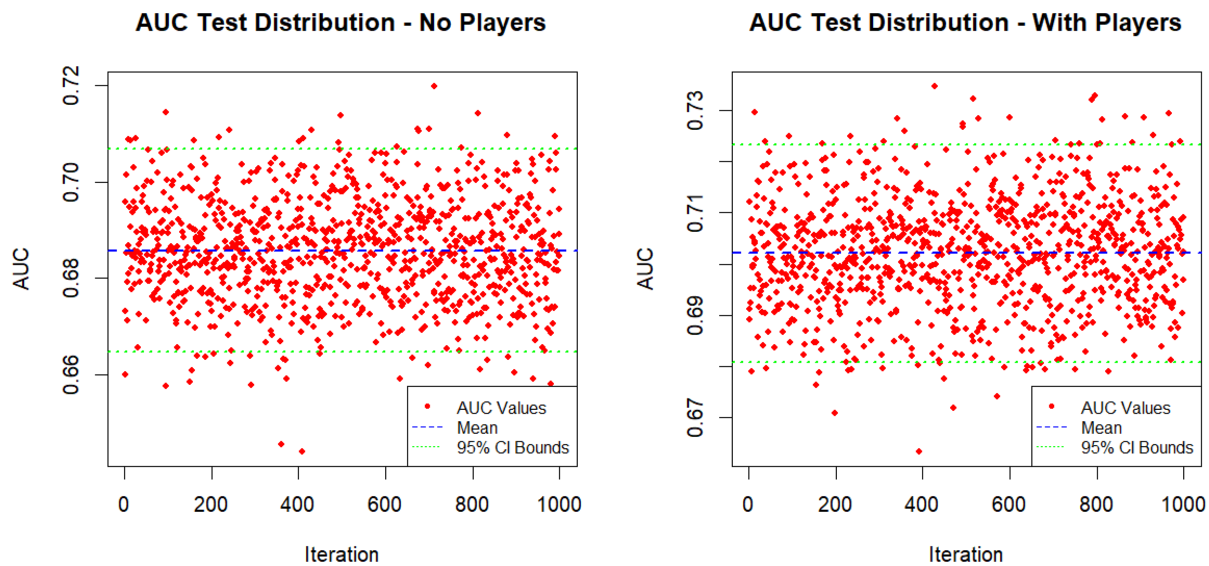
Our second explanation is that the starter measurements capture more about the culture, makeup and coaching of a team. If a team has "better" players they are more likely to convert a fourth down due to the skill of the players on the field but also from the experience that is passed onto other athletes. Similar to this there is more information about a team for the starter columns. This is because the "starters" of a team can be much more then just eleven players on offense and eleven players on defense. This is due to the variation in offensive but especially defensive formations.

5.3.4 Predictive Algorithms

In the following predictive analysis we have two main questions. First, are the conversions on fourth down predictable and do player specific factors help this process? Second do starter or on-field statistics contain more predictive power? The following figure displays the results of our XGBoosting algorithm in predicting the result of a fourth down attempt. The model was tested on 1000 bootstrapped test data sets after hyperparameters where tuned.

To answer our first question we create models that are run on two engineered datasets. The "No Players" data is the top 50 base columns, based on MDA, that have no player specific information. The “With Players” dataset then takes the top player variable at each position based on MDA and includes it along with our 50 base columns.

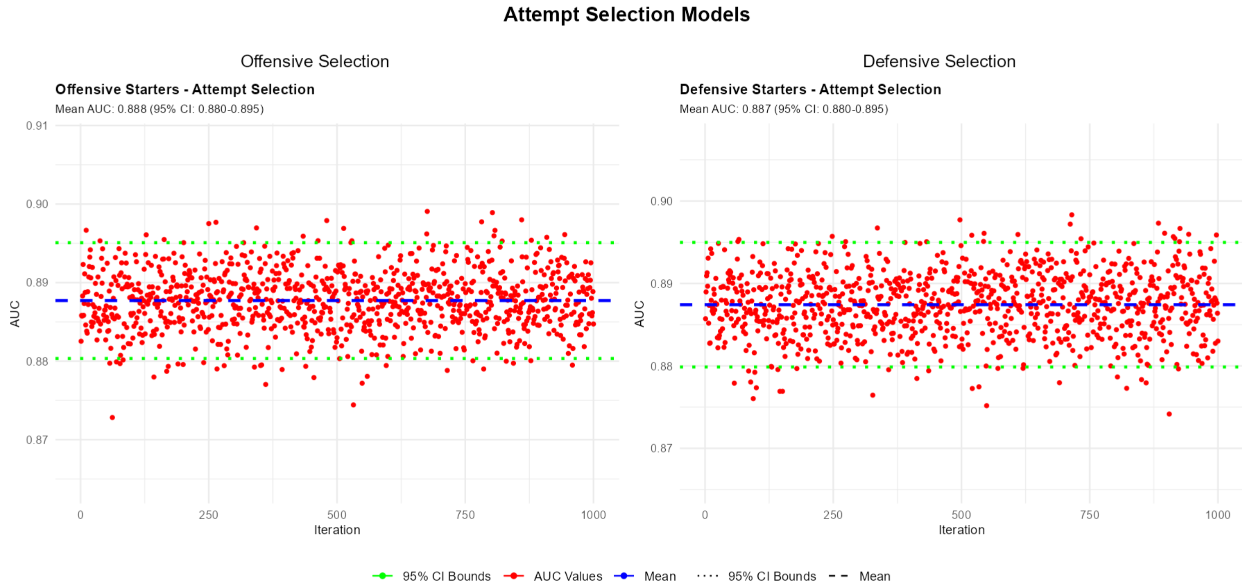
Figure 2: Predictive AUC



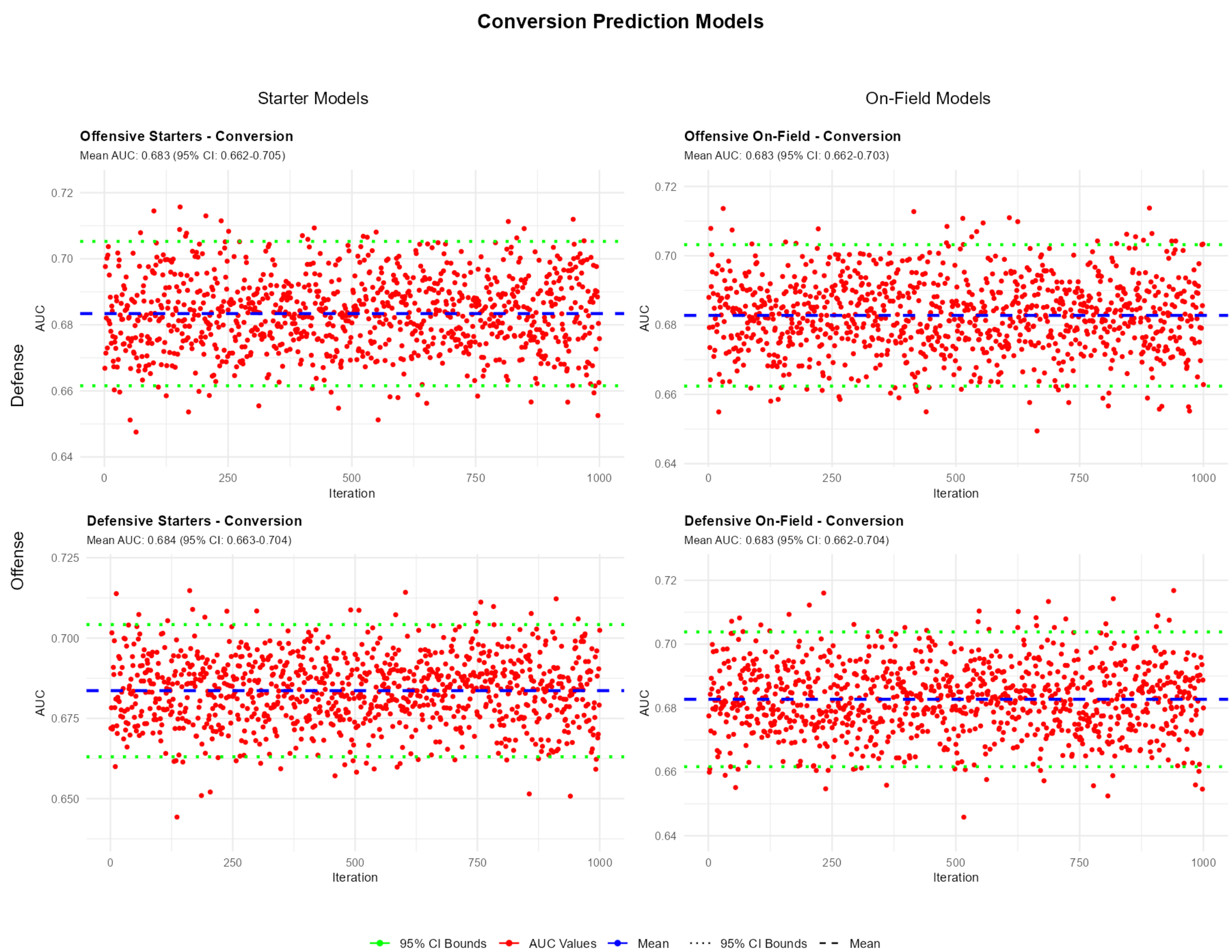
An AUC hovering around 0.7 is quite valuable when we consider the context of what we are predicting. What is most encouraging however is we see that AUC of models that possess player columns consistently perform better then the AUC of the datasets that do not contain player columns.

The following two sets of tables compare selection and outcome models when run with tuned XGBoosting on data that uses just on-field player statistics vs starter player statistics.

Table 11: Comparison of Selection and Conversion AUCs for different models



(a) Selection AUCs



(b) Conversion AUCs

From the previous tables we can reason that in terms of pure predictive power there is little difference in the AUC of the starter and on-field models. This highlights a need for further work as we are finding little to no difference in the predictive power of the two models but a difference in the statistical significance of the of the methods.

5.4 Discusion

In revisiting our original research questions we find encouraging but mixed answers to our questions. Our Generalized Inverse Mills Ratio is significant which both confirms selection bias (that we correct for) and tells us that coaches are able to factor in variables that we cannot see in our data. This is in line with theory, as we would imagine that coaches know their teams better than we do and therefore make better decisions. Our analysis of third down situations further supports this. It is found that kicker grades significantly impact third down play calling, particularly in the middle of the field where field goal attempts become viable if the third down fails. This demonstrates coaches' consideration of special teams abilities in their offensive decision-making, even before reaching fourth down. Despite this awareness, decision makers in the NFL are not perfect and fail to identify key archetypes of players such as Quarterbacks that excel in short passing. While a team can decide to simply adjust their decision making based on their own and opposing players, they also have the option of building a team to specifically excel in these situations. This would look like searching for Tight Ends that specifically excel in pass blocking or shifting the focus of the run game from the running back to the offensive line.

6 Future Research or Improvments

The biggest question that needs empirical answering is why the starter player measurements capture more statistical significance then the on-field measurement techniques. While some initial work has been done here, this is a scenario in which theory and robustness checks must be employed. Crowd attendance impact also can be further explored. We currently have relatively prepared data. This can be analyzed in different levels of coaches decision making as the 2020 pandemic provides a perfect situation to experiment within.

7 References

Back to Table of Contents

Cook, J. A. (2022). Sample-Selection-Adjusted Random Forests. International Journal of Data Science and Analytics, 14, 375-388. http://dx.doi.org/10.2139/ssrn.4225014

Das, M., Newey, W. K., & Vella, F. (2003). Nonparametric estimation of sample selection models. The Review of Economic Studies, 70(1), 33-58. https://doi.org/10.1111/1467-937X.00236

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology, 17(3), 295-314. https://doi.org/10.1016/0010-0285(85)90010-6

Goff, B. L., & Locke, C. (2019). Revisiting Romer: Digging Deeper Into Influences on NFL Managerial Decisions. Journal of Sports Economics, 20(5), 673-694. https://doi.org/10.1177/1527002518798686

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. Econometrica, 47(1), 153-161. https://doi.org/10.2307/1912352

Klein, R. W., & Spady, R. H. (1993). An efficient semiparametric estimator for binary response models. Econometrica, 61(2), 387-421. https://doi.org/10.2307/2951556

Lehman, D. W., & Hahn, J. (2013). Momentum and Organizational Risk Taking: Evidence from the National Football League. Management Science, 59(4), 852-868. https://doi.org/10.1287/mnsc.1120.1574

Losak, J. M., Stenquist, R., & Lovett, M. (2023). Behavioral Biases in Daily Fantasy

Baseball: The Case of the Hot Hand. Journal of Sports Economics, 24(3), 352-372. https://doi.org/10.1177/15270025221128955

Owens, M. F., & Roach, M. (2018). Decision-making on the hot seat and the short list: Evidence from college football fourth down decisions. Journal of Economic Behavior & Organization, 148, 226-245. https://doi.org/10.1016/j.jebo.2018.02.023

Roebber, P. J., Schultz, D. M., & Colle, B. A. (2022). On the existence of momentum in professional football. PLOS ONE, 17(6), e0269604. https://doi.org/10.1371/journal.pone.0269604

Romer, D. (2006). Do Firms Maximize? Evidence from Professional Football. Journal of Political Economy, 114(2), 340-365. https://doi.org/10.1086/501171

Yam, D., & Lopez, M. (2018). Quantifying the Causal Effects of Conservative Fourth Down Decision Making in the National Football League. Available at SSRN: https://ssrn.com/abstract=3114242 or http://dx.doi.org/10.2139/ssrn.3114242

Carl, S., Baldwin, B., Sharpe, L., Ho, T., Edwards, J. (2024). nflverse. Play-by-play data for NFL games. Retrieved from https://nflverse.nflverse.com/.

Pro Football Focus (PFF). (2024). NFL Player Performance Data. Accessed via subscription at https://premium.pff.com/nfl/teams/2024/REGPO.

Sporting News. (2024). Lions' Dan Campbell criticized for fourth-down decisions in playoff loss to 49ers. Retrieved from https://www.sportingnews.com/us/nfl/news/lions-49ers-dan-campbell-fourth-down-decisions/b54ce7cc05c4d55fc7b46285.

Davis, M. C., & End, C. M. (2010). A winning proposition: The economic impact of successful National Football League franchises. *Economic Inquiry*, 48(1), 39–50. https://doi.org/10.1111/j.1465-7295.2008.00124.x

Birkett, D. (2024, January 21). On this date: Memorable kneecap speech. Sports Illustrated. Retrieved from https://www.si.com/nfl/lions/news/on-this-date-memorable-kneecap-speech

Stine, R. (1989). An Introduction to Bootstrap Methods: Examples and Ideas. Sociological Methods & Research, 18(2-3), 243-291. https://doi.org/10.1177/0049124189018002003 (Original work published 1989)