

# A Predictive and Causal Analysis of Fourth Down Attempts in the NFL

## Context

The National Football League is a multi-billion dollar industry that has seen rapid growth within the American and International entertainment . 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 a age of data analytics in the NFL as teams seek to gain competitive advantages over their rivals. GMs and coaches then explore different avenues of strategies.

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. When a team reaches fourth down without achieving the needed yardage, they face a pivotal decision. Teams typically 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. 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 a 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.

## Research Problem and Glimpse of Findings

A result like this causes one 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. For example, 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 is that situation it could be argued that they should not have been as aggressive. This is because the panthers could have had a worse chance of being able to convert on fourth down.

On top of this, we must be weary 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 meeting. 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 GIMR that coaches are able to factor in variables that are unseen in our data in the decision making process on 4th downs. Coaches also show signs of being able to properly evaluate the performance of players in fourth down situations.

## 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, what key variables about players have predictive power in fourth down attempts? In other words, are there players that are more important in fourth down situations when compared to other situations. 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.

## Literature Review

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 or safe strategy in the NFL. He argues that teams value successful gambles more than the expected win percentage in a game. 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 aggressively than normal.

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

## Gap in Current Research

A gap in the current research is caused by the lack in 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 than 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.

## Roadmap

Go over this

## Data 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. The 2016 season however still had many inconsistencies which leads us to drop the year.

The key part then is the merging of Pro Football Focus's aggregated weekly data. The statistics are downloaded as premium player reports on a weekly bases and then aggregated to time lengths of 2 weeks, 12 weeks and 3 years. These time length have been arbitrarily picked which leaves room for future research. It was found that the 12 week timeframe better captured the significance and value of players. This is due to the 2 week timeframe struggling with noise while the 3 year timeframe is not sensitive enough to emerging or aging players.

Players from the PFF data set were merged into our base play by play data set based on a 12 step algorithm that matched ids across the two sources. The steps included pre-match ids, and then players matched on names, teams and positions.

As a final note attendance data was also scrapped from pro football reference.

## 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. In both situations the players are sorted into columns based on depth chart position. For example, the starting QB is also in the offense\_player\_1 slot to allow us to maintain consistency.

In our models to predict conversions in third or fourth downs we use the following base columns:

Our models that are used to predict the attempt of a fourth downs conversion then use the following base columns:

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 player that was called.

Statistic	Player Skill Metrics																																			
	QB Short	QB Med	QB Deep	RB Pass Bk	RB Run	WR1 Short Rt	WR1 Med Rt	WR1 Deep Rt	WR2 Run Bk	WR2 Short Rt	WR2 Med Rt	WR2 Deep Rt	WR3 Run Bk	WR3 Short Rt	WR3 Med Rt	WR3 Deep Rt	TE1 Pass Bk	TE1 Run Bk	TE1 Short Rt	TE1 Med Rt	TE1 Deep Rt	OL1 Pass Bk	OL1 Run Bk	OL2 Pass Bk	OL2 Run Bk	OL3 Pass Bk	OL3 Run Bk	OL4 Pass Bk	OL4 Run Bk	OL5 Pass Bk	OL5 Run Bk					
Min (ex. 0)	31	15	26	1	32	21	26	26	29	39	35	40	34	41	39	39	35	39	29	32	4	40	28	28	26	26	43	25	28	2	37	2	30	1	35	
Max	90	99	96	86	83	94	95	95	80	90	96	96	90	93	94	98	88	94	94	98	85	77	95	97	99	88	84	87	88	88	81	86	87	90	90	
Zero Count	164	209	220	297	175	290	2451	2800	139	174	201	237	665	730	778	843	1322	1495	1582	1713	229	159	253	473	1014	4	4	9	7	21	17	75	56	200	170	
5th Percentile	55	51	51	42	57	53	48	53	54	61	65	59	53	60	60	59	53	56	59	58	51	52	60	59	58	54	54	52	53	49	53	46	52	39	51	
Median	65	67	68	62	63	69	0	0	60	72	77	72	59	70	74	69	57	63	65	59	64	59	73	75	66	69	62	67	62	66	62	64	61	61	60	
Mean	62	63	64	57	61	64	27	20	58	69	72	68	50	58	61	57	40	43	44	41	59	57	69	66	54	68	62	66	62	65	62	62	60	58	58	
95th Percentile	72	79	82	77	71	81	90	91	67	82	86	86	67	81	86	89	67	80	86	91	72	67	85	91	95	79	72	79	72	78	72	76	71	76	71	
SD	14	17	18	20	14	20	35	32	12	16	18	19	23	28	30	30	28	34	36	36	16	12	19	26	33	8	6	9	6	10	7	13	9	18	14	
Starter Players																																				
Min (ex. 0)	34	24	27	4	32	26	26	26	35	40	29	40	40	40	35	40	36	26	35	39	9	39	29	28	28	13	32	1	37	2	28	2	30	16	38	
Max	86	94	96	84	79	94	95	95	82	92	95	96	87	91	95	95	90	92	95	98	82	74	94	96	99	87	88	88	89	88	89	88	86	87	85	
Zero Count	58	57	60	250	268	214	2424	2801	65	74	82	117	59	75	90	125	71	230	270	425	130	88	116	246	803	63	64	62	59	60	59	84	79	95	83	
5th Percentile	56	53	53	42	58	54	42	55	53	61	65	59	53	62	62	59	53	59	60	58	51	52	61	59	58	48	53	48	53	49	53	48	53	48	53	
Median	65	68	69	62	64	68	0	0	60	72	77	73	60	72	77	72	60	70	75	69	64	59	74	75	67	67	62	66	61	66	61	66	61	66	61	
Mean	64	66	67	57	60	65	27	20	59	71	75	71	74	70	74	70	59	66	69	63	61	58	71	71	57	65	61	64	61	64	61	64	61	63	60	
95th Percentile	72	78	81	76	72	81	90	91	67	83	86	86	67	82	86	85	67	80	86	90	72	67	85	91	93	79	72	78	72	78	71	78	72	77	72	
SD	9	11	12	18	17	18	34	32	9	11	13	15	8	11	13	15	9	18	20	24	13	10	14	20	31	13	10	12	9	13	11	13	10	13	10	
Select Data (Starter Only)																																				
Min (ex. 0)	34	24	27	4	32	26	26	26	35	40	29	40	33	40	35	40	34	26	35	39	9	39	29	28	28	13	32	1	37	2	28	2	30	16	32	
Max	86	94	96	84	79	94	95	95	83	92	97	96	87	91	95	96	90	93	95	98	82	74	94	100	99	87	88	88	89	88	89	88	86	87	88	
Zero Count	387	395	424	1811	1762	1474	11573	18398	388	452	563	772	374	482	546	780	431	1495	1805	2781	780	532	778	1566	5050	380	389	384	373	390	384	440	419	517	480	
5th Percentile	56	53	53	42	58	54	42	56	54	61	65	59	54	61	63	59	53	59	59	57	51	52	60	59	56	49	53	49	53	49	53	48	53	48	53	
Median	65	68	69	62	63	68	0	0	60	72	76	73	60	72	77	72	60	70	74	69	64	59	73	75	67	67	62	66	61	66	61	66	61	66	61	
Mean	64	66	67	57	60	64	27	19	59	71	75	70	71	74	70	59	66	68	63	61	58	71	71	57	65	61	64	61	64	61	64	61	63	60		
95th Percentile	72	79	81	76	71	82	90	91	67	82	86	86	67	81	86	85	67	80	86	91	72	67	85	91	91	79	72	78	72	78	71	78	72	77	71	
SD	9	11	12	19	17	18	35	31	8	11	13	15	8	12	13	15	9	18	20	24	13	9	15	20	30	12	10	9	12	9	13	10	13	10		

Figure 1: Player Offensive Summary

Statistic	Player Skill Metrics																																						
	DL1 Run Def	DL1 Pass Run	DL2 Run Def	DL2 Pass Run	DL3 Run Def	DL3 Pass Run	DL4 Run Def	DL4 Pass Run	LB1 Run Def	LB1 Man Cov	LB2 Zone Cov	LB2 Run Def	LB3 Man Cov	LB3 Zone Cov	LB4 Run Def	LB4 Man Cov	LB4 Zone Cov	CB1 Run Def	CB1 Man Cov	CB2 Zone Cov	CB2 Run Def	CB3 Zone Cov	CB3 Man Cov	CB4 Run Def	CB4 Zone Cov	CB5 Run Def	CB5 Man Cov	SL Run Def	SL Man Cov	SL Zone Cov	SL Run Def	SL Man Cov	SL Zone Cov						
Min (ex. 0)	42	45	35	49	28	49	32	47	27	33	29	26	28	28	25	32	32	27	44	29	42	28	33	33	26	30	26	28	40	47	29	38	30	34	39	43	29		
Max	83	86	84	89	84	90	96	81	79	86	82	82	80	80	88	93	88	88	89	91	75	96	93	79	96	88	82	92	91	76	82	90	74	82	90	74	85		
Zero Count	21	21	234	233	1065	1061	2053	2052	18	94	21	354	513	384	1706	1921	1759	2975	3133	3018	75	76	75	343	347	330	1585	1585	1573	36	42	43	482	499	483	3047	3049	3041	
5th Percentile	53	56	52	56	51	55	51	55	52	55	51	51	54	50	52	54	50	52	55	50	55	50	56	54	48	53	52	47	52	57	55	55	55	54	54	53	53		
Median	62	63	61	62	58	59	0	0	62	61	60	61	60	61	60	59	54	52	0	0	0	62	59	64	61	58	62	57	53	57	63	61	62	61	61	61	0	0	
Mean	62	63	58	60	45	46	29	30	62	60	60	56	53	54	35	32	33	16	13	15	61	58	62	56	54	58	37	35	38	62	61	61	54	53	54	15	15		
95th Percentile	72	74	71	75	69	73	68	69	71	69	68	70	68	68	69	68	68	69	68	68	68	71	68	69	71	68	69	71	69	67	68	68	68	68	69	67	67		
SD	7	7	15	16	27	28	30	31	7	10	7	18	21	18	30	31	30	27	25	26	9	10	10	10	18	18	18	30	29	31	7	7	8	20	21	20	26	26	
Starter Players																																							
Min (ex. 0)	30	51	29	45	43	49	26	48	26	32	28	29	35	29	26	33	28	26	28	28	38	27	31	27	27	29	33	28	26	41	43	30	38	29	45	29	34	30	
Max	80	88	81	89	80	83	83	93	90	93	87	81	86	79	82	88	79	80	98	80	82	78	81	82	87	95	77	89	90	75	75	92	79	82	86	77	79	90	
Zero Count	60	60	54	54	53	53	109	109	61	171	77	55	161	82	135	243	164	333	807	538	69	70	72	74	73	72	220	261	215	68	76	77	67	91	79	538	660	544	
5th Percentile	53	56	52	56	52	56	51	55	52	55	50	52	55	51	51	55	50	51	54	49	54	49	54	49	54	53	46	50	54	53	46	54	55	54	54	53	52	52	
Median	62	62	62	62	62	62	60	61	61	61	59	61	61	60	61	61	59	60	60	59	62	59	63	61	58	63	60	58	62	61	61	62	61	61	60	60	60		
Mean	61	62	61	62	61	62	59	60	60	59	58	60	59	58	59	57	57	55	49	51	60	58	62	60	57	61	57	54	58	61	60	60	61	59	60	52	50	52	
95th Percentile	71	73	71	73	71	74	70	73	70	73	70	68	67	71	68	67	70	68	68	69	68	67	68	68	71	68	68	71	68	68	71	69	67	68	69	67	66	67	
SD	9	9	9	9	9	9	11	11	9	13	10	9	13	10	12	15	13	17	25	21	9	10	10	10	9	10	10	15	16	15	9	9	10	9	10	10	21	23	21
Select Data (Starter Only)																																							
Min (ex. 0)	30	50	29	45	39	49	26	48	26	32	28	27	35	29	26	27	28	26	28	28	38	24	36	31	27	26	29	33	28	26	41	43	30	38	29	29	29	26	
Max	80	88	83	80	80	83	83	93	90	93	91	81	89	79	82	88	79	84	98	80	82	84	93	81	82	87	95	82	89	90	75	87	92	79	82	86	77	79	90
Zero Count	353	353	336	336	330	330	712	713	734	1088	508	584	1100	521	827	1686	998	2100	5058	3484	405	425	431	431	442	1462	1852	1470	1148	457	457	467	612	563	3080	3943	3173		
5th Percentile	53	56	53	56	52	56	52	55	52	54	50	52	55	51	54	50	51	49	54	49	54	49	54	49	54	53	46	50	54	53	46	54	55	56	54	53	52	52	
Median	62	62	62	62	62	62	61	60	61	61	59	61	61	60	61	61	60	60	59	62	59	63	62	59	63	62	59	63	60	58	61	62	61	61	62	61	61	60	60
Mean	61	62	61	62	61	62	59	60	60	59	58	60	59	58	59	57	57	55	49	51	61	58	62	60	58	62	57	54	58	61	60	60	61	60	60	53	51	53	
95th Percentile	71	73	71	74	71	73	70	72	71	68	68	71	68	68	70	68	68	69	67	68	68	71	68	68	71	68	68	71	69	67	68	68	69	68	68	67	66	68	
SD	9	9	9	9	9	9	11	12	9	13	10	9	13	10	12	16	13	17	25	21	9	10	10	10	9	10	10	15	16	16	9	9	9	9	10	10	22	22	

Statistic	Game Situation Metrics									
	Conversion	Attempt	Yards To Go	Temperature	Wind	Vegas Win Prob	Spread Line	Total Line	Punter Grade	Kicker Grade
Min	0	Not in out-come	1.0	6.0	0.0	0.0	-18.0	30.0	Not in out-come	Not in out-come
Max	1		34.0	93.0	44.0	1.0	22.0	63.5		
Mean	0.52		4.1	60.8	5.3	0.3	2.0	45.4		
Median	1		2.0	70.0	4.0	0.2	3.0	45.5		
SD	0.5		4.6	15.3	5.9	0.3	6.3	4.6		
Zero Count	1923		0.0	0.0	1604.0	0.0	0.0	0.0		
5th Percentile	0		1.0	33.0	0.0	0.0	-7.5	38.0		
95th Percentile	1		14.0	80.0	16.0	1.0	13.0	53.5		
SELECT_DI										
Select Data										
Min	Not in select	0	1.0	6.0	0.0	0.0	-18.0	30.0	0	0
Max		1	46.0	97.0	44.0	1.0	22.0	63.5	78.9	86.1
Mean		0.2	7.8	61.7	5.4	0.5	1.8	45.0	60.4	59.7
Median		0	7.0	70.0	4.0	0.4	2.5	45.0	61.2	61.8
SD		0.4	5.7	15.1	5.7	0.3	6.4	4.6	7.7	11.6
Zero Count		19970	0.0	0.0	9639.0	0.0	44.0	0.0	376	825
5th Percentile		0	1.0	34.0	0.0	0.0	-8.0	38.0	57.5	51.9
95th Percentile		1	19.0	82.0	16.0	1.0	13.0	53.0	65	67.4

Figure 3: Other Variable Summary

## Methodology and 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/Team 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.

Tools that have both predictive and causal traits are being employed.

In our data we often had 20,000+ different player variables. This clearly is a situation where predictive tools such as Random Forest and XGBoosting thrive. In the use of XGBoosting we first use bootstrapping to tune for optimal hyperparameters. Second we will run the tuned model on 1000 train/test bootstrapped splits of data to report the AUC. This process will be done when predicting attempts and conversion.

Similarly we can enact Random Forest to receive a OOB 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.

For the causal analysis we will be employing the 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 a evolved version 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 is 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 a attempt of a fourth down. This prediction then is converted into a Generalized IMR that is not bound to the assumptions of a IMR. We then place our GIMR in our outcome equation when estimating the conversion of a fourth down to control for selection bias.

### Step 1: Selection Model

Model the binary choice to attempt a play:

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

The probability is modeled as:

$$z_i = \text{RF}(\mathbf{X}_i, \mathbf{K}_i, \mathbf{P}_i) \quad (2)$$

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 generalized inverse Mills ratio is:

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

The first and second order conditions for  $\lambda_i$  are:

$$\begin{aligned} \text{FOC: } \frac{\partial \lambda_i}{\partial z_i} &= \frac{1}{(1 - z_i)^2} > 0 \text{ for } z_i < 1 \\ \text{SOC: } \frac{\partial^2 \lambda_i}{\partial z_i^2} &= \frac{2}{(1 - z_i)^3} > 0 \text{ for } z_i < 1 \end{aligned}$$

Thus  $\lambda_i$  is strictly increasing and convex in  $z_i$ .

## Step 2: Outcome Model

Model the binary conversion outcome for attempted plays:

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

Linear probability model with selection correction:

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

where:

- $\mathbf{X}_i$ : covariates excluding kicker and punter grades
- $\lambda_i$ : generalized inverse Mills ratio
- $\boldsymbol{\beta}$ : coefficient vector for main covariates
- $\theta$ : selection correction parameter
- $\varepsilon_i$ : 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).

## Findings

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 luxury 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 3rd down to be treated similarly to a fourth down by the 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 3rd and 10, the coach will treat this attempt as a “last chance”.

We can first note that the grades of a teams kicker is positively significant in it’s effect on the decision to attempt a fourth down. Exogeneity 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. However 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 apart of the decision making process on third down, similarly to fourth down, it does not have a effect on 3rd down conversion.

Variable	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
<b>Variables of Interest:</b>										
Punter Grades (12w)	-3.02***	-0.57	-1.71*	-4.25***	-1.13	-1.91*	-0.09	-0.21	0.10	0.37
Kicker FG Grades (12w)	1.66*	-0.77	0.65	3.21***	1.85*	2.79***	1.64.	0.42	0.65	0.49
<b>Sample Size</b>	<b>988</b>	<b>3,023</b>	<b>6,865</b>	<b>6,924</b>	<b>5,805</b>	<b>4,945</b>	<b>4,271</b>	<b>3,897</b>	<b>3,478</b>	<b>3,080</b>
<b>Controls:</b>										
Game Situation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Environmental	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	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Season/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coach FE	No	No	No	No	No	No	No	No	No	No

*Note:* Values shown are t-statistics from OLS regressions with conversion as the dependent variable

Significance codes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1, . p<0.15

Figure 4: Kicker Offensive Grades

This significance of the kicker and punter grades in key area 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 kicker are and making adjustments on their playcalling.



## Offensive Grades Model Results

Variable	Starter Models			On-Field Models			Attempt Models		
	Starter LPM	Starter Probit	Starter Logit	On-Field LPM	On-Field Probit	On-Field Logit	Attempt LPM	Attempt Probit	Attempt Logit
<b>Quarterbacks</b>									
Short Grades Pass	0.029(0.016)*	0.090(0.049)*	0.148(0.082)*	0.031(0.014)**	0.097(0.044)**	0.165(0.072)**	0.007(0.004)	0.030(0.021)	0.052(0.038)
Medium Grades Pass	-0.015(0.013)	-0.043(0.041)	-0.073(0.068)	-0.002(0.015)	-0.000(0.045)	-0.009(0.074)	0.006(0.004)	0.022(0.017)	0.038(0.031)
Deep Grades Pass	-0.013(0.013)	-0.043(0.040)	-0.075(0.066)	-0.021(0.013)	-0.072(0.041)*	-0.117(0.068)*	0.006(0.004)	0.021(0.017)	0.046(0.030)
<b>Running Backs</b>									
Grades Pass Block	-0.021(0.010)**	-0.065(0.031)**	-0.108(0.053)**	0.012(0.010)	0.034(0.030)	0.057(0.050)	-0.004(0.003)	-0.024(0.014)*	-0.041(0.024)*
Grades Run	0.012(0.011)	0.038(0.032)	0.065(0.054)	-0.004(0.011)	-0.014(0.033)	-0.025(0.055)	-0.001(0.003)	-0.004(0.013)	-0.002(0.023)
Short Grades Pass	0.004(0.011)	0.013(0.031)	0.018(0.052)	-0.003(0.010)	-0.006(0.031)	-0.012(0.051)	-0.000(0.003)	0.001(0.014)	0.001(0.024)
Medium Grades Pass	-0.015(0.009)	-0.044(0.028)	-0.072(0.047)	-0.005(0.009)	-0.015(0.027)	-0.025(0.044)	-0.001(0.003)	-0.006(0.013)	-0.009(0.022)
Deep Grades Pass	-0.005(0.010)	-0.013(0.029)	-0.024(0.048)	-0.003(0.009)	-0.007(0.027)	-0.014(0.045)	0.003(0.003)	0.008(0.012)	0.014(0.021)
<b>WR1</b>									
Grades Run Block	0.006(0.015)	0.014(0.046)	0.030(0.077)	-0.011(0.019)	-0.037(0.058)	-0.054(0.098)	-0.006(0.004)	-0.024(0.019)	-0.046(0.034)
Short Grades Pass	0.003(0.013)	0.008(0.041)	0.020(0.068)	0.000(0.016)	-0.004(0.046)	-0.009(0.077)	-0.000(0.004)	-0.001(0.017)	-0.007(0.030)
Medium Grades Pass	0.015(0.013)	0.044(0.040)	0.075(0.066)	0.000(0.014)	0.003(0.043)	0.005(0.071)	0.005(0.004)	0.015(0.018)	0.034(0.031)
Deep Grades Pass	0.001(0.011)	0.001(0.034)	0.002(0.056)	0.001(0.012)	0.011(0.036)	0.017(0.059)	0.003(0.003)	0.016(0.014)	0.027(0.025)
<b>WR2</b>									
Grades Run Block	-0.021(0.016)	-0.062(0.047)	-0.101(0.077)	0.028(0.036)	0.087(0.111)	0.152(0.181)	-0.001(0.004)	-0.003(0.020)	0.005(0.035)
Short Grades Pass	-0.010(0.013)	-0.031(0.039)	-0.056(0.066)	0.014(0.024)	0.041(0.072)	0.067(0.120)	0.008(0.004)**	0.036(0.016)**	0.059(0.029)**
Medium Grades Pass	0.043(0.035)	0.075(0.058)	0.104(0.090)	-0.004(0.020)	-0.020(0.060)	-0.029(0.101)	0.001(0.004)	-0.000(0.016)	-0.001(0.027)
Deep Grades Pass	0.003(0.011)	0.014(0.033)	0.020(0.054)	0.010(0.016)	0.041(0.049)	0.062(0.082)	0.000(0.003)	0.002(0.013)	0.003(0.024)
<b>WR3</b>									
Grades Run Block	-0.007(0.014)	-0.021(0.042)	-0.041(0.068)	0.017(0.027)	0.060(0.082)	0.093(0.137)	-0.001(0.004)	-0.002(0.017)	-0.004(0.031)
Short Grades Pass	-0.015(0.011)	-0.048(0.033)	-0.076(0.055)	-0.002(0.020)	-0.002(0.061)	0.009(0.102)	-0.003(0.003)	-0.014(0.014)	-0.021(0.025)
Medium Grades Pass	0.007(0.010)	0.021(0.031)	0.036(0.051)	0.011(0.019)	0.017(0.057)	0.029(0.094)	0.006(0.003)*	0.021(0.014)	0.039(0.024)
Deep Grades Pass	0.003(0.010)	0.013(0.030)	0.017(0.049)	0.015(0.015)	0.056(0.046)	0.087(0.076)	0.001(0.003)	0.009(0.013)	0.014(0.022)
<b>TE1</b>									
Grades Pass Block	0.026(0.012)**	0.077(0.035)**	0.132(0.058)**	0.024(0.011)**	0.079(0.034)**	0.130(0.058)**	-0.002(0.003)	-0.012(0.014)	-0.016(0.026)
Grades Run Block	-0.015(0.015)	-0.044(0.045)	-0.077(0.074)	-0.012(0.012)	-0.040(0.038)	-0.067(0.063)	-0.000(0.004)	0.007(0.019)	0.007(0.034)
Short Grades Pass	0.004(0.013)	0.009(0.037)	0.014(0.063)	0.000(0.013)	-0.003(0.038)	-0.006(0.063)	0.003(0.004)	0.010(0.018)	0.021(0.030)
Medium Grades Pass	-0.016(0.011)	-0.049(0.032)	-0.082(0.053)	0.017(0.011)	0.051(0.032)	0.088(0.052)*	0.002(0.003)	0.008(0.014)	0.008(0.024)
Deep Grades Pass	-0.004(0.010)	-0.012(0.029)	-0.018(0.049)	-0.015(0.009)	-0.046(0.028)	-0.076(0.047)	0.002(0.003)	0.009(0.012)	0.018(0.022)
<b>OL1</b>									
Grades Pass Block	0.016(0.012)	0.053(0.036)	0.090(0.060)	-0.010(0.010)	-0.030(0.029)	-0.050(0.048)	-0.004(0.003)	-0.017(0.015)	-0.027(0.027)
Grades Run Block	0.004(0.014)	0.015(0.042)	0.019(0.070)	0.006(0.010)	0.018(0.030)	0.035(0.050)	0.001(0.004)	0.000(0.017)	0.002(0.031)
<b>OL2</b>									
Grades Pass Block	-0.009(0.012)	-0.025(0.035)	-0.042(0.058)	-0.005(0.010)	-0.015(0.029)	-0.026(0.048)	-0.000(0.003)	-0.006(0.015)	-0.015(0.026)
Grades Run Block	-0.003(0.014)	-0.009(0.042)	-0.014(0.070)	-0.012(0.010)	-0.038(0.030)	-0.068(0.049)	0.003(0.004)	0.020(0.017)	0.037(0.031)
<b>OL3</b>									
Grades Pass Block	0.009(0.012)	0.026(0.035)	0.050(0.059)	0.007(0.009)	0.024(0.029)	0.039(0.048)	0.001(0.003)	0.006(0.015)	0.006(0.027)
Grades Run Block	-0.015(0.014)	-0.047(0.043)	-0.080(0.070)	0.004(0.009)	0.016(0.029)	0.023(0.048)	-0.004(0.004)	-0.019(0.018)	-0.038(0.032)
<b>OL4</b>									
Grades Pass Block	0.022(0.012)*	0.065(0.036)*	0.112(0.060)*	0.007(0.010)	0.018(0.030)	0.033(0.050)	-0.003(0.003)	-0.016(0.015)	-0.023(0.026)
Grades Run Block	-0.006(0.013)	-0.021(0.039)	-0.038(0.065)	-0.013(0.010)	-0.036(0.029)	-0.063(0.048)	0.000(0.004)	0.005(0.017)	0.007(0.030)
<b>OL5</b>									
Grades Pass Block	0.031(0.012)**	0.089(0.037)**	0.149(0.062)**	-0.012(0.010)	-0.038(0.031)	-0.063(0.051)	0.001(0.003)	0.015(0.015)	0.020(0.027)
Grades Run Block	-0.031(0.013)**	-0.093(0.040)**	-0.151(0.067)**	0.010(0.010)	0.026(0.030)	0.043(0.051)	-0.002(0.004)	-0.010(0.017)	-0.020(0.030)
<b>Perf. Measures</b>									
GIMR	0.009(0.002)***	0.025(0.007)***	0.041(0.011)***	0.008(0.002)***	0.025(0.007)***	0.040(0.011)***			
<b>Control Variables</b>									
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/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Coach FE	No	No	No	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

Figure 5: Offense Grades Table

## Defensive Grades Model Results

Variable	Starter Models			On-Field Models			Attempt Models		
	Starter LPM	Starter Probit	Starter Logit	On-Field LPM	On-Field Probit	On-Field Logit	Attempt LPM	Attempt Probit	Attempt Logit
<b>DL1</b>									
Grades Run Defense	0.001(0.015)	0.005(0.046)	0.006(0.077)	-0.005(0.011)	-0.014(0.034)	-0.020(0.056)	-0.006(0.004)	-0.023(0.019)	-0.035(0.034)
Grds Pass Rush Def	0.009(0.015)	0.030(0.045)	0.053(0.074)	-0.014(0.011)	-0.045(0.034)	-0.077(0.056)	0.008(0.004)*	0.036(0.019)*	0.068(0.033)**
<b>DL2</b>									
Grades Run Defense	0.026(0.015)*	0.079(0.044)*	0.128(0.073)*	0.005(0.024)	0.022(0.071)	0.031(0.119)	-0.002(0.004)	-0.009(0.019)	-0.010(0.034)
Grds Pass Rush Def	-0.017(0.014)	-0.051(0.042)	-0.088(0.070)	-0.018(0.023)	-0.059(0.069)	-0.093(0.116)	-0.003(0.004)	-0.012(0.019)	-0.032(0.033)
<b>DL3</b>									
Grades Run Defense	-0.005(0.015)	-0.017(0.045)	-0.028(0.075)	-0.015(0.042)	-0.051(0.123)	-0.076(0.202)	-0.004(0.004)	-0.028(0.020)	-0.048(0.035)
Grds Pass Rush Def	0.003(0.015)	0.015(0.044)	0.020(0.073)	-0.001(0.045)	0.011(0.134)	-0.009(0.220)	0.001(0.004)	0.016(0.019)	0.024(0.033)
<b>DL4</b>									
Grades Run Defense	0.012(0.015)	0.031(0.044)	0.055(0.073)	-0.042(0.050)	-0.136(0.151)	-0.227(0.252)	0.000(0.004)	-0.010(0.020)	-0.007(0.035)
Grds Pass Rush Def	-0.004(0.015)	-0.008(0.044)	-0.017(0.073)	-0.017(0.053)	-0.040(0.157)	-0.079(0.262)	-0.002(0.004)	0.001(0.021)	-0.007(0.035)
<b>LB1</b>									
Grades Run Defense	0.027(0.013)**	0.084(0.039)**	0.139(0.065)**	0.005(0.010)	0.014(0.030)	0.023(0.050)	0.001(0.004)	0.001(0.017)	-0.003(0.030)
Man Grades Cov Def	-0.022(0.009)**	-0.069(0.029)**	-0.110(0.048)**	0.005(0.008)	0.018(0.025)	0.028(0.042)	0.002(0.003)	0.013(0.013)	0.028(0.023)
Zone Grades Cov Def	-0.007(0.012)	-0.020(0.037)	-0.030(0.062)	-0.002(0.010)	-0.010(0.030)	-0.012(0.049)	-0.005(0.004)	-0.018(0.015)	-0.039(0.027)
<b>LB2</b>									
Grades Run Defense	0.026(0.013)**	0.084(0.040)**	0.137(0.067)**	-0.014(0.017)	-0.040(0.052)	-0.067(0.086)	0.002(0.004)	0.018(0.017)	0.030(0.031)
Man Grades Cov Def	-0.011(0.010)	-0.033(0.030)	-0.057(0.051)	0.004(0.013)	0.015(0.040)	0.020(0.067)	-0.001(0.003)	-0.007(0.013)	-0.009(0.023)
Zone Grades Cov Def	-0.014(0.012)	-0.043(0.035)	-0.069(0.058)	0.013(0.017)	0.033(0.051)	0.057(0.084)	0.002(0.003)	0.013(0.015)	0.021(0.026)
<b>LB3</b>									
Grades Run Defense	-0.004(0.013)	-0.012(0.040)	-0.022(0.067)	-0.036(0.036)	-0.099(0.109)	-0.173(0.178)	-0.002(0.004)	-0.011(0.016)	-0.023(0.029)
Man Grades Cov Def	0.018(0.011)	0.053(0.033)	0.089(0.056)	0.018(0.019)	0.054(0.055)	0.091(0.092)	0.005(0.003)*	0.018(0.014)	0.038(0.024)
Zone Grades Cov Def	-0.001(0.012)	-0.005(0.038)	-0.004(0.064)	-0.042(0.028)	-0.124(0.084)	-0.211(0.140)	0.005(0.004)	0.027(0.016)*	0.046(0.028)*
<b>LB4</b>									
Grades Run Defense	-0.029(0.012)**	-0.084(0.035)**	-0.146(0.058)**	0.032(0.040)	0.084(0.119)	0.142(0.199)	0.002(0.003)	0.010(0.016)	0.017(0.027)
Man Grades Cov Def	0.024(0.012)**	0.072(0.035)**	0.120(0.058)**	-0.003(0.018)	-0.016(0.054)	-0.025(0.090)	-0.000(0.003)	-0.004(0.015)	-0.001(0.026)
Zone Grades Cov Def	0.003(0.013)	0.011(0.040)	0.021(0.065)	-0.003(0.029)	-0.005(0.091)	-0.012(0.152)	-0.001(0.004)	-0.007(0.016)	-0.017(0.029)
<b>CB1</b>									
Grades Run Defense	-0.013(0.015)	-0.039(0.045)	-0.059(0.075)	0.008(0.016)	0.027(0.049)	0.042(0.081)	-0.002(0.004)	-0.009(0.019)	-0.018(0.034)
Man Grades Cov Def	-0.010(0.013)	-0.034(0.038)	-0.052(0.064)	-0.029(0.014)**	-0.083(0.042)**	-0.144(0.069)**	0.000(0.004)	0.002(0.017)	0.011(0.029)
Zone Grades Cov Def	0.007(0.015)	0.025(0.044)	0.036(0.073)	-0.012(0.015)	-0.036(0.046)	-0.056(0.077)	0.000(0.004)	-0.003(0.018)	-0.012(0.032)
<b>CB2</b>									
Grades Run Defense	0.033(0.014)**	0.104(0.042)**	0.176(0.070)**	-0.019(0.020)	-0.054(0.060)	-0.093(0.098)	-0.008(0.004)*	-0.023(0.018)	-0.048(0.032)
Man Grades Cov Def	-0.013(0.012)	-0.046(0.036)	-0.072(0.060)	0.008(0.018)	0.023(0.055)	0.036(0.091)	-0.004(0.004)	-0.016(0.016)	-0.024(0.027)
Zone Grades Cov Def	-0.011(0.014)	-0.037(0.042)	-0.062(0.070)	-0.007(0.021)	-0.023(0.064)	-0.035(0.106)	0.002(0.004)	0.002(0.018)	0.008(0.031)
<b>CB3</b>									
Grades Run Defense	-0.018(0.012)	-0.060(0.036)	-0.094(0.061)	0.032(0.033)	0.100(0.098)	0.160(0.163)	-0.005(0.004)	-0.027(0.016)*	-0.050(0.029)*
Man Grades Cov Def	0.018(0.012)	0.062(0.037)*	0.101(0.062)	-0.001(0.030)	0.009(0.087)	0.017(0.144)	0.011(0.004)**	0.043(0.016)**	0.077(0.029)**
Zone Grades Cov Def	-0.010(0.013)	-0.028(0.041)	-0.052(0.069)	0.030(0.034)	0.087(0.100)	0.145(0.167)	0.002(0.004)	0.016(0.018)	0.031(0.031)
<b>S1</b>									
Grades Run Defense	0.002(0.016)	0.008(0.048)	0.019(0.081)	-0.020(0.013)	-0.058(0.040)	-0.097(0.067)	-0.003(0.005)	-0.016(0.023)	-0.021(0.039)
Man Grades Cov Def	-0.006(0.014)	-0.014(0.043)	-0.028(0.071)	0.000(0.012)	-0.004(0.039)	-0.008(0.066)	-0.003(0.004)	-0.020(0.019)	-0.040(0.033)
Zone Grades Cov Def	-0.006(0.014)	-0.021(0.044)	-0.034(0.073)	0.013(0.012)	0.044(0.037)	0.071(0.063)	0.006(0.004)	0.027(0.019)	0.048(0.033)
<b>S2</b>									
Grades Run Defense	-0.008(0.015)	-0.026(0.047)	-0.044(0.079)	-0.002(0.027)	-0.009(0.078)	-0.024(0.129)	-0.002(0.005)	-0.017(0.020)	-0.023(0.036)
Man Grades Cov Def	-0.009(0.013)	-0.024(0.040)	-0.038(0.066)	-0.008(0.025)	-0.026(0.079)	-0.041(0.131)	0.003(0.004)	0.014(0.018)	0.015(0.033)
Zone Grades Cov Def	-0.004(0.014)	-0.015(0.041)	-0.025(0.068)	0.006(0.026)	0.024(0.080)	0.042(0.134)	0.002(0.005)	0.010(0.020)	0.024(0.036)
<b>S3</b>									
Grades Run Defense	-0.007(0.013)	-0.018(0.040)	-0.029(0.067)	0.009(0.044)	0.022(0.139)	0.027(0.237)	-0.002(0.004)	-0.005(0.016)	-0.001(0.028)
Man Grades Cov Def	-0.003(0.013)	-0.010(0.040)	-0.018(0.066)	-0.018(0.043)	-0.041(0.127)	-0.070(0.213)	0.004(0.004)	0.012(0.016)	0.026(0.029)
Zone Grades Cov Def	0.020(0.015)	0.062(0.045)	0.102(0.075)	0.019(0.047)	0.046(0.146)	0.089(0.247)	-0.005(0.004)	-0.018(0.018)	-0.039(0.032)
<b>Perf. Measures</b>									
GIMR	0.007(0.002)***	0.024(0.007)***	0.042(0.012)***	0.007(0.002)***	0.024(0.007)***	0.040(0.012)***			
<b>Control Variables</b>									
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/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Coach FE	No	No	No	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

Figure 6: Defense Grades Table

## GIMR

In examining the following tables the 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.

## Players

The exogeneity of kicker grades then 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 or defensive players. If a 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 it’s ability to capture the overall performance of a player in a specific area. The other measures will be kept as robustness checks.

To evaluate a 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.

### Offensive Player Joint Significance

Variable	Starter Models			On-Field Models			Attempt Models		
	Starter LPM	Starter Probit	Starter Logit	On-Field LPM	On-Field Probit	On-Field Logit	Attempt LPM	Attempt Probit	Attempt Logit
Quarterbacks	1.80 [3,4016]	1.88 [3,4016]	1.91 [3,4016]	2.40* [3,4016]	2.72** [3,4016]	2.74** [3,4016]	2.64** [3,25605]	1.74 [3,25605]	1.95 [3,25605]
Running Backs	1.68 [5,4014]	1.71 [5,4014]	1.69 [5,4014]	0.41 [5,4014]	0.38 [5,4014]	0.39 [5,4014]	0.65 [5,25603]	0.76 [5,25603]	0.71 [5,25603]
Wide Receivers	0.68 [12,4007]	0.70 [12,4007]	0.72 [12,4007]	0.29 [12,4007]	0.37 [12,4007]	0.33 [12,4007]	1.11 [12,25596]	1.04 [12,25596]	1.03 [12,25596]
Tight Ends	1.72 [5,4014]	1.71 [5,4014]	1.79 [5,4014]	2.12* [5,4014]	2.35** [5,4014]	2.34** [5,4014]	0.35 [5,25603]	0.39 [5,25603]	0.35 [5,25603]
Offensive Line	1.98** [10,4009]	1.92** [10,4009]	1.97** [10,4009]	0.82 [10,4009]	0.85 [10,4009]	0.91 [10,4009]	0.42 [10,25598]	0.66 [10,25598]	0.63 [10,25598]
<b>Control Variables</b>									
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/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Coach FE	No	No	No	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. Stars indicate significance of joint test for all variables in position group.

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

Figure 7: Offense Grades Model Results

### Defensive Player Joint Significance

#### Offensive Players

Quarterbacks have been undoubtedly held as the most important player on the field. Signs of joint significance are found across Starter, On-Field and Attempt models. On-Field models show clear significance while the Attempt model also shows signs of statistical significance.

For individual Quarterback Feature importance we find that the Quarterbacks performance on short passes is significant. Since this is held across all models it shows us that coaches are able to properly evaluate the performance of a quarterback in a fourth down situation. However, one difference is found with the On-Field deep passing grades. The negative significance suggests that quarterbacks that are better at deep passing are less likely to be effective on a fourth down. Coaches are not picking up on this feature as there is no negative significance in the Attempt Models.

Variable	Starter Models			On-Field Models			Attempt Models		
	Starter LPM	Starter Probit	Starter Logit	On-Field LPM	On-Field Probit	On-Field Logit	Attempt LPM	Attempt Probit	Attempt Logit
Defensive Line	0.74 [8,4011]	0.75 [8,4011]	0.75 [8,4011]	0.44 [8,4011]	0.48 [8,4011]	0.47 [8,4011]	0.90 [8,25600]	1.10 [8,25600]	1.10 [8,25600]
Linebackers	2.44*** [12,4007]	2.51*** [12,4007]	2.52*** [12,4007]	0.59 [12,4007]	0.54 [12,4007]	0.55 [12,4007]	0.75 [12,25596]	0.86 [12,25596]	0.99 [12,25596]
Cornerbacks	1.55 [9,4010]	1.81* [9,4010]	1.74* [9,4010]	0.91 [9,4010]	0.88 [9,4010]	0.89 [9,4010]	1.84* [9,25599]	1.47 [9,25599]	1.62 [9,25599]
Safeties	0.36 [9,4010]	0.37 [9,4010]	0.37 [9,4010]	0.42 [9,4010]	0.44 [9,4010]	0.43 [9,4010]	0.81 [9,25599]	0.76 [9,25599]	0.80 [9,25599]
<b>Control Variables</b>									
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/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Coach FE	No	No	No	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. Stars indicate significance of joint test for all variables in position group.  
\* \* p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01

Figure 8: Defense Grades Model Results by Position Groups

One big case of coaches misevaluating a player is found by how a Tight Ends pass blocking grades is consitantly significant across On-Field and Starter Models. However in our Attempts model we find no such significance.

Coaches are including wide receivers in the decision making process as seen by their joint significance in the Attempt Models. However, there is a struggle to find consistant significance in the individual models especially with the On-Field and Starter models.

What is the most interesting is that the Grades of the team's running back are not found to be significant in any model. This is the case even despite how fourth downs often are run plays due to the amount of short yardage situations.

## Defensive Players

In terms of the defense there is not enough evidence to make statements about individual player features' significance. However, when we test for joint significance of the defensive players there is signs of significance with the Defensive Line and Corner backs in which the the coaches respond to.

## Offense Joint Significance

Dataset	Starter Models			On-Field Models			Attempt Models		
	Starter LPM	Starter Probit	Starter Logit	On-Field LPM	On-Field Probit	On-Field Logit	Attempt LPM	Attempt Probit	Attempt Logit
Offense Grades	1.44*** [35,3984]	1.44** [35,3984]	1.47** [35,3984]	0.90 [35,3984]	0.99 [35,3984]	1.00 [35,3984]	0.87 [35,25573]	0.86 [35,25573]	0.85 [35,25573]
Offense Yards	1.97*** [35,3984]	2.06*** [35,3984]	2.05*** [35,3984]	1.05 [35,3984]	1.07 [35,3984]	1.06 [35,3984]	0.80 [35,25573]	0.96 [35,25573]	0.88 [35,25573]
Offense Completions	1.70*** [35,3984]	1.76*** [35,3984]	1.74*** [35,3984]	1.19 [35,3984]	1.20 [35,3984]	1.18 [35,3984]	0.69 [35,25573]	0.80 [35,25573]	0.71 [35,25573]
<b>Control Variables</b>									
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/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Coach FE	No	No	No	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	No	No	No	No	No	No	No	No	No
Control Off. Players	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*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

Figure 9: Joint Significance Offense

## Defense Joint Significance

Dataset	Starter Models			On-Field Models			Attempt Models		
	Starter LPM	Starter Probit	Starter Logit	On-Field LPM	On-Field Probit	On-Field Logit	Attempt LPM	Attempt Probit	Attempt Logit
Defense Grades	1.38* [38,3981]	1.47** [38,3981]	1.45** [38,3981]	0.59 [38,3981]	0.58 [38,3981]	0.59 [38,3981]	1.05 [38,25570]	1.03 [38,25570]	1.12 [38,25570]
Defense Stops	1.31* [38,3981]	1.33* [38,3981]	1.31* [38,3981]	0.84 [38,3981]	0.80 [38,3981]	0.83 [38,3981]	1.17 [38,25570]	1.12 [38,25570]	1.09 [38,25570]
Defense Tackles	1.13 [38,3981]	1.11 [38,3981]	1.10 [38,3981]	0.79 [38,3981]	0.80 [38,3981]	0.78 [38,3981]	1.12 [38,25570]	1.07 [38,25570]	1.06 [38,25570]
<b>Control Variables</b>									
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/Team FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Coach FE	No	No	No	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

Figure 10: Joint Significance Defense

## References

- 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](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>.

Raymond, S. (2024). *4th Down Analysis App*. Interactive dashboard for analyzing 4th down decisions in the NFL. Available at: [https://jzmtko-yigit-aydede.shinyapps.io/NFL\\_4th\\_Down\\_App/](https://jzmtko-yigit-aydede.shinyapps.io/NFL_4th_Down_App/).

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>