To Return or Not to Return?

A statistical model to determine whether to return a kick in American Football

Reynald Oliveria and Chris Desrochers

Executive Summary

We analyze the decision problem in American Football of when to return kicks received in the endzone and when to take a touchback. Using the PROACTURD approach for decision problems, we first define the problem, objectives, and alternatives. Our audience desires to receive the best score differential possible on each drive resulting from a kick, and they choose between returning the kick in an attempt to start their drive past the twenty-five yard line and taking a touchback and starting their drive at the twenty-five yard line. Next, we summarize the two models we combined to predict the expected points of the resulting drive of a kick based upon the return location where the kick was received and the distance the ball was kicked during the play. These models are trained on a comprehensive dataset of NFL plays from 2009 to 2019 (Yurko), and their effectiveness in describing the uncertainty in return success is determined.

We consider the tradeoffs and consequences of each alternative, to aid in interpreting the models. We discuss the outcome differences between kick returns and touchbacks, considering the levels of uncertainty, the potential gains and losses, and the status of the touchback as a sort of "control" alternative without negative consequences. Risk is considered in this discussion as the alternatives have largely different risks, given a touchback prevents a play from happening entirely. We also describe how each individual decision to return a kick or take a touchback affects future decisions. Based upon these factors and the results of the models themselves, we conclude first that neither alternative is Pareto dominated. Specifically we find that kicks should be returned when the receiving team is losing by at least a touchdown or has the last possession of a half, and that touchbacks should be taken otherwise. We discuss how this general strategy will change based upon the individual situation of each kick (including considerations of game momentum, risk, and necessity of gain).

Introduction

In American Football, coaches and players often must make quick decisions based upon their observations of the game scenario. One of these kinds of decisions takes place during kicks. A kick is a play where the kicking team either just scored or ran out of chances to progress down the field during an offensive possession. The kicking team gives possession of the ball to the receiving team, who picks a kick receiver to stand, ready to catch and return the kick. If a kick lands in the endzone or is caught by the receiver and he takes a knee in the endzone (or a few other less important cases), the kick results in a touchback and the receiving team starts the drive at their own 25-yard line.

Kicks can either be punts or kickoffs. Punts are when the kicking team is giving up possession after running out of chances to progress down the field. Punts can be kicked from anywhere on the field, depending on where the kicking team is on the field. Kickoffs occur when the kicking team just scored and at the start of each half. Kickoffs start from the kicking team's 30-yard line, and often the ball is kicked into the endzone.

For our final project, we analyzed the decision problem of when to take a touchback and when to return the ball when receiving a kick in the endzone in American Football. The objective of American Football is to score more points than the other team, but for our problem the objective is more specific. When receiving a kick, a team wants to maximize the point differential of the resulting drive. This maximization is subject to other situational factors; teams may not have time to drive all the way down the field, or teams may be leading the other team and may simply want to prevent the other team from scoring. Thus, the objective for determining what action to take with a kick is to provide the best chance of scoring.

During a kick, the receiver can choose to return the kick, let the kick bounce, call a fair catch (to have the offense start their drive where he catches the ball), or not interact with the ball. However, for our problem, the alternatives can be narrowed down to returning the kick and taking a touchback. This is because in the endzone, a fair catch is the same as taking a knee, which results in a touchback. In addition, letting the ball bounce (or not interacting with it) in the endzone also results in a touchback. Thus, the alternatives are to return the kick or to take a touchback.

Our Models

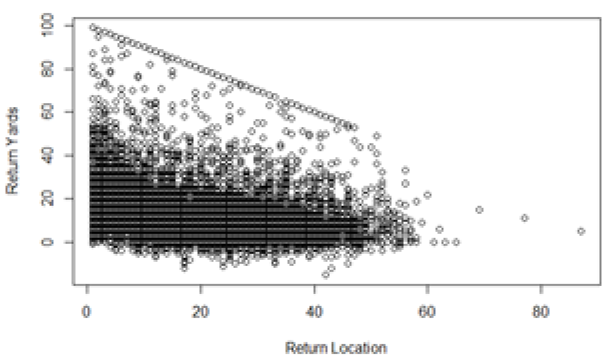
Overview. To inform the decision of whether to return a kick or take a touchback, we created two models. The first model takes in the context in which one is receiving a kick (data about the kick) and predicts the number of return yards the receiving team would gain if they choose to return the kick. This result, using the location from which the return started, can easily be converted to another metric which is the number of yards until the team who received the ball needs to drive before reaching their opponents endzone. This metric to which the result of the first model is converted will be known as the *yards-to-go* of the team in possession. While more return yards, or consequently less yards-to-go, always mean a better outlook for the team, it is not clear from this model alone the value of return yards. For example, it is not obvious whether the proportion of the values of 10 yards-to-go and 20 yards-to-go should be different from the proportion of the values of 70 yards-to-go and 80 yards-to-go. And so, to solve this problem, we created a second model that takes in yards-to-go on a first down and predicts the change in score of the next scoring play from the perspective of the team in possession. Because these models are not perfect, uncertainty is embedded in their predictions. This uncertainty from the models is analogous to the uncertainty that players and coaches face concerning the consequences of their decisions. Therefore, we interpret the results of the models as distributions rather than singular predicted values. Finally, these two models were created using a play-by-play dataset (Yurko) that contains data about all the plays of all NFL games (pre-season, regular season, post-season) from 2009-2019.

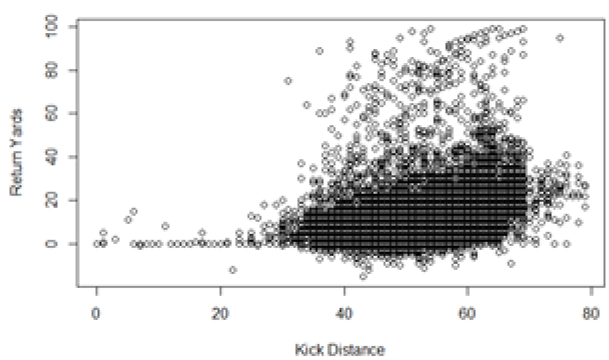
Return Yards Model. We first want to create a model that predicts expected return yards for a given kick. During a kick, the kicking team sends rushing players running down the field at the receiving team's returner while blockers from the receiving team attempt to slow them down. Though each of these blocker/rusher matchups are non-measurable influencers on the return distance, there are still measurable parts of the kick itself that account for some of the variance in return distance. For instance, the time the ball is in the air has a notable impact upon the return distance. If the ball is in the air longer, the rushers have more time to progress down the field at the returner before a return can be started, since the ball has not been caught yet.

However, our dataset does not have a hang time statistic. Instead of hang time, we chose to use kick distance, which is the distance from the line of scrimmage (where the ball was kicked from) to the yard line where the ball was caught. Looking at the statistics present in our dataset, we chose to build the model based upon this kick distance and a metric we called "return location." This metric was calculated based upon a few different statistics in the dataset and is a number from 0 to 100 representing the location on the field where the return started (where 0 is the returning team's own goal line).

Before building a model, we filtered the dataset to include only kick plays. Additionally, we filtered out plays where there were fumbles, fair catches, or unreturnable kicks. Our model is intended to simply predict the return distance from a successful return, so it should not be trained on plays where a successful return was not made. In our problem, a successful return is one where the receiver caught the ball and attempted a return. It is still possible to have negative yardage on such a return, so the return is only necessarily successful in that the returner tried to return the ball.

With a quick visual inspection of these two variables and their relationships to the number of return yards, we decide to calculate a linear regression across this limited data where the number of return yards is modeled by a weighted linear sum of the position of the return and the kick distance. Based upon the graphs, shown below, we did not have concrete evidence that a higher degree polynomial model was necessary. The resulting regression model is also shown.





$$Return\ distance\ =\ -\ 0.121\ \times\ Return\ Location\ +\ 0.474\ \times\ Kick\ Distance\ -\ 8.922$$

This model has an R^2 of 18.5%. This means that the variance of these two factors explain 18.5% of the variance of the return distance. This is not incredibly significant but is still useful for generally predicting return distance. This significance level is not unexpected, because of the previous discussion of the many non-quantitative factors that affect this distance. Our model can still be used to have a general idea of how far a kick can be returned, and the other variance in the return distance is noted for future reflection.

We can also affirm the coefficients learned by the model, as it is intuitive that a higher return location (the ball is received closer to the kicking team) would result in a lower return distance, and a higher kick distance would result in a higher return distance. This can be seen in the negative coefficient for the return location and a positive coefficient for the kick distance. In addition, the magnitude of the return location coefficient is lower than the magnitude of the kick distance coefficient, which seems accurate since the kick distance would have a large influence on the return distance. The intercept coefficient is not as meaningful as the other coefficients, as the return distance when the kick distance and return location are both zero is not meaningful. This would be the case where a kick was made from the returning team's goal line to the returning team's goal line. However, considering kickoffs, when the return location is zero and the kick distance is seventy, the return distance is just under twenty-five, which makes sense considering a touchback is twenty-five yards.

Next Score. Given a situation in a game, the *next score* is essentially the change in score of the next scoring play from the perspective of the team in possession. While technically, a try for either an extra point or a two-point conversion after a touchdown is a different play from that which resulted in the touchdown, we will consider the tries after a touchdown to be the same play as that which resulted in the touchdown. For example, suppose the Jets and the Giants were in a game. The Jets are on a first down with 20 yards-to-go. This situation will be referred to as the "origin situation." We will then calculate two next scores depending on two different sequences of events that occur from this origin situation.

a) The Jets complete a 10-yard pass earning them a first down with 10-yards to go. They then run 6 yards, earning them a second down with 4-yards to go. Finally, they run and make a touchdown. They also attempt a kick for an extra-point, and it was good. The Jets have scored a total of 7 points.

b) The Jets attempt a 10-yard pass, and the Giants intercept and the Giants player with the ball goes down on the 35-yard line on the Jets' half, earning the Giants a first down with 35 yards-to-go. The Giants then complete a 35-yard pass earning a touchdown. They attempt and fail a two-point conversion. The Giants have scored a total of 6 points.

In (a), the scoring play that most immediately succeeds the origin situation earns the Jets 7 points. Since the Jets are the team in possession in the origin situation, the next score of the origin situation given (a) is +7 points. In (b), the scoring play that most immediately succeeds the origin situation earns the Giants 6 points. Since the Jets, not the Giants, are the team in possession, the next score of the origin situation given (b) is -6 points.

Expected Points of a Touchdown. As discussed, after a touchdown, the scoring team can attempt to score more points. The scoring team can either try to score an extra point with a kick or go for a two-point conversion. Because we consider this extra try to be part of the play, then a scoring play resulting in a touchdown has an expected point value of more than 6 points. By looking at all the touchdowns and the tries thereafter available in the dataset, we can calculate that the expected point value of a touchdown is 6.9 points with a variance of 0.1521.

Expected Points Model. We ran a linear regression between the yards-to-go on a first down and the next score with an intercept at 6.9. The intercept is at 6.9 as 0 yards-to-go implies a touchdown, which as discussed, has an expected point value of 6.9. The predicted next score from this model can therefore be interpreted as the expected points of the play that results in a situation with a first down and the inputted yards-to-go. We assumed a linear relationship as a previous study (Burke), calculating expected value of each possible value of yards-to-go rather than performing a regression, results in a graph that looks like a linear relationship. The regression results in this equation:

$$Next\ score\ on\ a\ First\ Down = 6.9 - 0.0808 \times Yards\ to\ Go$$

The model predicts that for each yard further the team in possession away from their goal endzone, the expected points decreases by 0.0808. This makes sense as it is harder to score the further you are from your goal endzone.

Predictions Using the Models. To predict the expected points from a touchback, one would need to input 75 yards to go into the Expected Points Model as a touchback is a guaranteed 25 yards. To convert this into a distribution, we use the expected points as the mean. In this case, the expected points are 0.84, then we calculate a variance. The variance of the touchback expected point distribution is the sum of the Mean Squared Error of the Expected Points Model added to the Variance of a touchdown. We treat Mean Squared Error like variance as variance is the Mean Squared Error of a model with 0 predictors. And with that, the variance of expected points from a touchback is predicted to be 23.91.

Similarly, this model can generate a distribution of expected points when one chooses to return a kick. First, the distance of the kick must be inputted into the Return Yards model. For the touchback decision to be relevant, the return location must be 0 yards from the receiving team's endzone, and the kick distance must equal that of how many yards away the kick happens from the receiving team's endzone. And so, one only needs to input the distance in yards between from where the kick happens and the receiving team's endzone to get the proper result from the Return Yards model. The output will be used as the mean, and the variance of the result will be the Mean Squared Error of the Return Yards model which is 137.12. We then plug in this distribution to the Expected Points Model; the mean is calculated as in the formula above. The variance is calculated assuming independence and using variance addition rules like so:

$$Var(Expected\ Points) = 0.0808^2 \times MSE(Return\ Yards) + Var(Touchdown) + MSE(Next\ Score)$$

The Mean Squared Error of the Expected Points model is 23.76, and so the variance of expected points from a return is 24.80. With the means and variances of expected points from a touchback and a return, we can then compare probabilities of the next score being n points or more, or $-n$ points or less. We use a normal distribution to find these probabilities.

Score Difference									
-8	-7	-6	-3	-2	2	3	6	7	8
7.16%	10.39%	14.57%	32.94%	40.63%	28.06%	21.61%	8.09%	5.44%	3.53%

The table above depicts the probability that after the next scoring play, the losing team will no longer be losing based on the score difference if the team receiving the kick decides to go for a touchback. For example, if the receiving team is losing by 3, then the probability of them either tying up the score or winning at the next score is 32.94%. On the other hand, if the receiving team is winning by 7 points, then the probability that the score will be tied, or they will be losing at the next score is 5.44%. And so, when the score difference is negative, the receiving team hopes for higher probabilities. And if the score difference is positive, the receiving team hopes for lower probabilities. The score differences chosen are all the possible next scores (National Football League).

	Score Difference									
Kick Yards	-8	-7	-6	-3	-2	2	3	6	7	8
100	11.14%	15.43%	20.69%	41.49%	49.43%	21.51%	16.11%	5.57%	3.65%	2.31%
99	11.00%	15.25%	20.47%	41.19%	49.13%	21.73%	16.30%	5.65%	3.71%	2.35%
98	10.85%	15.07%	20.25%	40.89%	48.82%	21.96%	16.49%	5.74%	3.77%	2.39%
97	10.71%	14.89%	20.03%	40.59%	48.51%	22.19%	16.68%	5.83%	3.84%	2.44%

96	10.57%	14.71%	19.82%	40.29%	48.21%	22.42%	16.88%	5.92%	3.90%	2.48%
95	10.43%	14.53%	19.61%	40.00%	47.90%	22.65%	17.07%	6.01%	3.97%	2.53%
94	10.29%	14.36%	19.39%	39.70%	47.60%	22.88%	17.27%	6.10%	4.03%	2.57%
93	10.16%	14.19%	19.18%	39.40%	47.29%	23.11%	17.46%	6.20%	4.10%	2.62%
92	10.02%	14.01%	18.97%	39.11%	46.98%	23.35%	17.66%	6.29%	4.17%	2.67%
91	9.88%	13.84%	18.77%	38.81%	46.68%	23.58%	17.86%	6.39%	4.24%	2.71%
90	9.75%	13.68%	18.56%	38.52%	46.37%	23.82%	18.07%	6.49%	4.31%	2.76%
89	9.62%	13.51%	18.35%	38.23%	46.07%	24.06%	18.27%	6.58%	4.38%	2.81%
88	9.49%	13.34%	18.15%	37.93%	45.76%	24.30%	18.47%	6.68%	4.45%	2.86%
87	9.36%	13.18%	17.95%	37.64%	45.46%	24.54%	18.68%	6.78%	4.52%	2.91%
86	9.23%	13.01%	17.75%	37.35%	45.15%	24.78%	18.89%	6.88%	4.60%	2.96%
85	9.11%	12.85%	17.55%	37.06%	44.85%	25.03%	19.09%	6.99%	4.67%	3.02%
84	8.98%	12.69%	17.35%	36.77%	44.54%	25.27%	19.30%	7.09%	4.75%	3.07%
83	8.86%	12.53%	17.15%	36.48%	44.24%	25.52%	19.52%	7.19%	4.82%	3.12%
82	8.73%	12.37%	16.96%	36.19%	43.94%	25.77%	19.73%	7.30%	4.90%	3.18%
81	8.61%	12.22%	16.77%	35.90%	43.63%	26.01%	19.94%	7.41%	4.98%	3.23%
80	8.49%	12.06%	16.57%	35.62%	43.33%	26.26%	20.16%	7.52%	5.06%	3.29%
79	8.37%	11.91%	16.38%	35.33%	43.03%	26.52%	20.38%	7.63%	5.14%	3.35%
78	8.26%	11.76%	16.19%	35.04%	42.73%	26.77%	20.59%	7.74%	5.22%	3.40%
77	8.14%	11.60%	16.01%	34.76%	42.42%	27.02%	20.81%	7.85%	5.30%	3.46%
76	8.02%	11.45%	15.82%	34.48%	42.12%	27.28%	21.03%	7.96%	5.39%	3.52%
75	7.91%	11.31%	15.63%	34.19%	41.82%	27.53%	21.26%	8.08%	5.47%	3.58%
74	7.80%	11.16%	15.45%	33.91%	41.52%	27.79%	21.48%	8.19%	5.56%	3.64%
73	7.69%	11.01%	15.27%	33.63%	41.22%	28.05%	21.71%	8.31%	5.64%	3.70%
72	7.58%	10.87%	15.09%	33.35%	40.92%	28.31%	21.93%	8.43%	5.73%	3.77%
71	7.47%	10.73%	14.91%	33.07%	40.63%	28.57%	22.16%	8.55%	5.82%	3.83%
70	7.36%	10.59%	14.73%	32.79%	40.33%	28.83%	22.39%	8.67%	5.91%	3.89%
69	7.25%	10.45%	14.55%	32.51%	40.03%	29.10%	22.62%	8.79%	6.00%	3.96%

The table above depicts the probability that after the next scoring play, the losing team will no longer be losing based on the score difference if the team receiving the kick decides to go for a return. For example, if the receiving team is losing by 3, and they are receiving a kick that reaches the endzone from 80 yards away, then the probability of them either tying up the score or winning at the next score is 35.62%. On the other hand, if the receiving team is winning by 7 points, and they are receiving a kick that reaches the endzone from 71 yards away, then the probability that the score will be tied, or they will be losing at the next score is 5.82%. As discussed earlier, when the score difference is negative, the receiving team hopes for higher probabilities. And if the score difference is positive, the receiving team hopes for lower probabilities. Hence, the probabilities highlighted in green in the table above, represents the situations in which there is a better probability that the receiving team is not losing at the next score if they choose to return instead of taking the touchback. The other possible kick yards not depicted in the table above have less favorable probabilities than that of the touchback. And so, from the table above, policy can be made to optimize the probability of not losing at the next score. A policy for kickoffs (70 kick yards), for example, is that the receiving team should only return a kickoff if they are losing by a touchdown.

Conclusions and Decision Strategy

Based upon our models, we make a decision strategy for our problem. Before that, we must consider the factors that define this decision strategy.

There are numerous tradeoffs between the two alternatives. Choosing the return provides a chance for more than twenty-five yards, yet it provides less consistency than a touchback. A touchback prevents the possibility of a loss of possession or a return of less than twenty-five yards, but the receiving team also does not have an opportunity for a momentum-building play through the return.

Similarly, there are tangible consequences to account for in the decision strategy. The return guarantees more uncertainty than the touchback, leading to an outcome that is harder to predict. The return can result in negative results including injured players or loss of possession, yet it also could result in psychological momentum or a penalty on the kicking team. Conversely, a touchback generally has no consequences (positive or negative) since it acts as a sort of control option. It has tradeoffs compared to other options, but inherently it does not have consequences other than the twenty-five-yard gain, since no play occurs.

The importance of defining a decision strategy can be seen most clearly through the risks in the decision, which are related to the tradeoffs and the consequences. When returning a kick, the receiving team could lose possession of the ball or receive less than twenty-five yards. Additionally, since a play occurs when taking a return, there is a risk of injury. Touchbacks are an option without risk, since the twenty-five-yard gain cannot be a negative result unless it is in comparison to another option. These risk factors are important to note because their importance will not only be dependent on the decision strategy, but the scenario of the decision.

Our decision strategy will also need to consider various forms of linked decisions relevant to this problem. Though each decision to return a kick or take a touchback does not change the actual problem or objectives of the next decision, each choice will affect the context of the next choice.

Given the mean number of return yards from a kick return and from a touchback, choosing to always take a touchback will theoretically bring you more yards throughout the game. Since the mean number of return yards is lower (as predicted from our model) than the yards gained by a touchback, as samples are drawn from the distribution of kick returns, the mean number of return yards will approach the mean of the distribution because of the law of large numbers.

Taking these factors into account, we define our decision strategy without a specific utility function or risk tolerance in mind. The specifics of the decision strategy would be dependent on each coach, but we define general strategies that would be most effective.

Coaches should generally take a touchback when they have an opportunity, as it is a safer option. They can avoid many risks altogether, and still gain a significant number of yards compared to the number they would gain from a return. However, coaches should consider the possibility of taking a return in a few cases. If their team is down by a single score with a low amount of time in a half, they should most likely take a return. This gives them an opportunity to quickly progress down the field.

Additionally, if the receiving team is receiving a kickoff, they should return the kick any time they are down by at least a touchdown. The expected chance of tying the score or getting a higher score was found to be higher for seventy-yard kicks when the score differential was a touchdown in the kicking team's favor.

The decision on when to return a kick or take a touchback is fluid. We find that generally teams want to take a return when the chance of benefit is necessary to make a difference in the game. This condition is dependent upon coaches' risk profiles, but we find that there is reason to take a return when the receiving team is down by at least a touchdown, or when there is only time for one final possession in a half.

Bibliography

Burke, B. (2014, December 07). Retrieved November/December, 2020, from https://www.advancedfootballanalytics.com/index.php/home/stats/stats-explained/expected-points-and-epa-explained

National Football League. (n.d.). Scoring Plays: NFL Football Operations. Retrieved November/December, 2020, from https://operations.nfl.com/the-rules/nfl-video-rulebook/scoring-plays/

Yurko, R. (n.d.). Ryurko/nflscrapR-data. Retrieved November/December, 2020, from https://github.com/ryurko/nflscrapR-data/tree/master/play_by_play_data