# Analysis of NFL Two-Point Conversions

Yusuf Mian, Hamzeh Hamdan

April 2025

## 1 Introduction and Motivation

NFL coaches often take major flak from the public over their decision making on two-point attempts in games. After every six-point touchdown, a coach must decide to take a two-point attempt to score from the two yard line or to kick a shot one point extra-point attempt. Take a December 2023 game between the Detroit Lions and Dallas Cowboys where Lions Coach Dan Campbell opted to go for a two-point conversion while losing by one point in an attempt to win the game rather than tie the game. His team went on to lose the game and fans were split on whether he was responsible for the loss. Lots of the decision making comes down to game situation, fatigue, game plan, and even kicker strength.

Given the increased importance of analytics in football, the goal of our project is to try to understand if we can model when coaches go for two-point attempts over one-point extra-point attempts and to see if we can model how success changes within a game. The vast majority of attempts are extra-points rather than two-point attempts, but our goal is to isolate the situations that call for two-point attempts and consider how strategy can be maximized in those situations.

Some of these are more obvious situations - like a team using a two-point conversion to tie the game or make it a one score game as opposed to two. Others are more recent trends, such as the analytics based argument for teams to go for a two-point conversion when down by 8 points in order to make it a 6 point game. The rationale behind this argument is that when a team is down by 7 points (which they would be if they kick an extra point down 8) and then ties the game with another touchdown, they still only have a 50-50 shot at best of winning the game. If a team converts on the two-point conversion and then scores a touchdown, they can kick an extra point to actually win the game. And if they fail, they still have another opportunity to tie the game when down 8 with a two-point conversion on the next touchdown. We will both explore if this is a strategy that teams are using and if this is something that they should be using.

**Table 1: Attempt Counts**

| Attempt Type | Count | Percentage |
|---|---|---|
| Extra-point | 12082 | 95.1 |
| Two-point | 616 | 4.9 |

## 2 Data and Exploratory Data Analysis

We will be using NFL play-by-play data obtained from the `nflfastR` package, a package that contains advanced NFL data and play-by-play data. The package's play-by-play data extends from 1999 to 2024 (the most recent NFL season). The data mostly tracks the same variables across each season with a couple of exceptions. These variables include hundreds of details on the specifics of the play (the game date, the quarter, the time, the offense and defense, the down, the yards to go, the yard line, the play type, a description of the play, indicator variables for certain types of plays like — pass, rush, completed pass, touchdown, sack, fumble, interception, field goal, two-point conversion, etc.). It also contains some advanced analytics like probabilities of certain outcomes happening, WPA, EPA, etc.

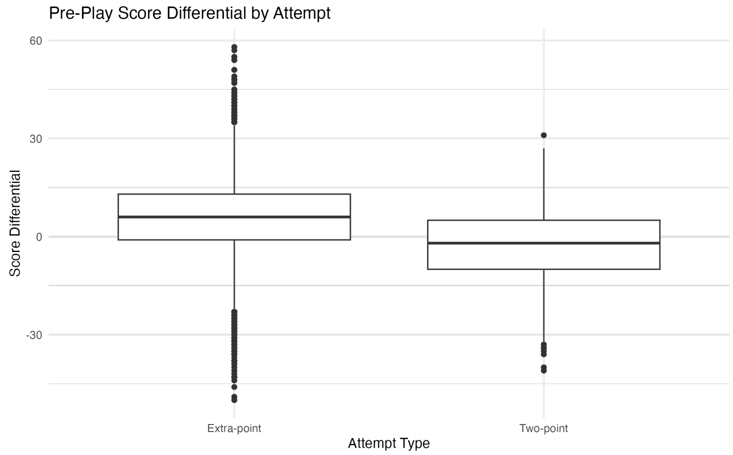
For our purposes, the most relevant variables are some binary indicators (`two_point_attempt` and one we will make from turning `two_point_conv_result` into a binary indicator) as well as some of the details of the play like `qtr`, `time`, `posteam` (team with possession), `defteam` (team on defense), `posteam_score` (score of the possession team before the play), `defteam_score`, `posteam_score_post` (score of the possession team after the play), `defteam_score_post`, `score_differential_post`, `two_point_conversion_prob` (probability of making a two-point conversion), `extra_point_prob`, and potentially later adding some more covariates to our model.

The EDA analyses were broken into three types: figuring out when in the game coaches choose between a two-point try and an extra-point kick, how the score and win-probability context drives the choice, and which teams are more open to the 2-point attempt.

We explored 2001 to 2024 data, looking at distributions by quarter, half, and minutes remaining. We analyzed how the score differential and total points correlate to the aggressiveness chosen by the play.

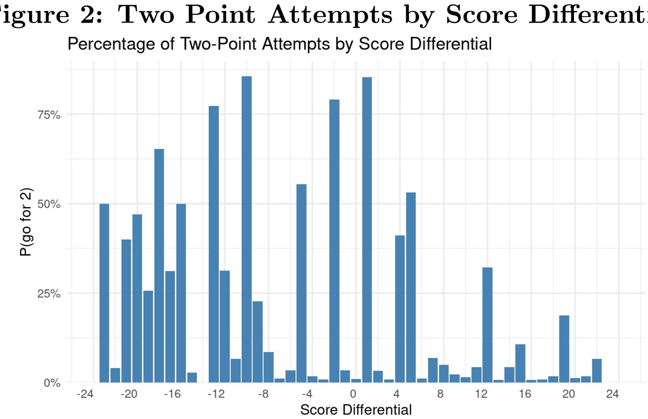
One of the first thing we analyzed, seen in Figure 1, is how frequent each type of play was chosen and how this changed as the score differential changes. We saw that the extra-point was far more likely to occur, which matches our intuition from the game. We also saw that the extra point kick had a sharper distribution and had a larger mean of score differential than the two-point attempt. This makes sense; teams that are winning are less incentivized to take the risk of the two-point. On the other hand, teams that are down are much more likely to chose the two-point attempt.

**Figure 1: Pre-Play Score Differential by Attempt**




Since scores are discrete and not continuous, we also considered how likely teams were to go for two at different score differentials to identify specific scenarios. As Figure 2 shows, some of the highest peaks of going for two occur when a team is down 13, down 10, down 2, up 1, and up 5. All of these scenarios intuitively make sense. When a team is down 13, they have the opportunity to change the deficit from 12, which requires two touchdowns, to 11, which requires a touchdown and a field goal. When a team is down 10, they can make the deficit 8, which is only one touchdown, by going for 2. When a team is down 2, they obviously will go for a two point conversion to tie the game. When a team is up 1 or 5, they will go for a two point conversion to increase the lead to either a 7 point game or 3 point game.

**Figure 2: Two Point Attempts by Score Differential**




The only scenario which surprised us in Figure 2 is the relatively low density at a differential of -8. As we discussed in the motivation, we believe this to be a potentially interesting scenario where teams could increase their win probability by going for two. We will return to this scenario in our results exploration.

**Figure 3: Two-Point Attempt Rate Over Time**

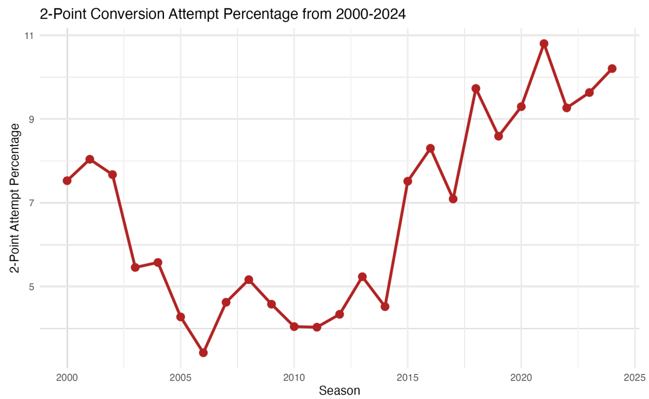
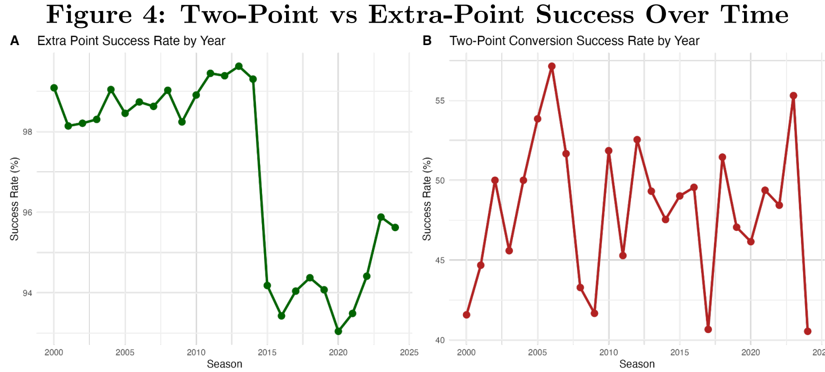



**Figure 4: Two-Point vs Extra-Point Success Over Time**




We also explored how the scoring play percentages changed over time, and we saw a recent increase in the two-point attempts. We were rather interested in this, especially the jump from 2013/2014 to 2015/2016. After doing some research, we found that this was when the NFL moved the line of scrimmage for extra-point kicks from the 2-yard line back to the 15-yard line. This also marked a lean towards analytics to re-evaluate the expected points of each option. As extra point success decreased due to the rule change, there was a push to consider more two point conversions and this follows an increase in attempts. Due to this trend, we will only consider the years post-rule change (2015-present) in our analysis.

## 3 Methods

Our methods involve two larger forms of analysis. First, we will try to model how and when coaches decide to go for two-point conversions. For this, we will use logistic regression to predict decisions to go for two using all the play-by-play data where it is possible to go for a conversion (all plays post-touchdown which are either an extra point or two-point conversion). We fit three different models.

**Table 2: Coaches 2-Point Attempt Percentages**

Top & Bottom 5 Coaches by Two-Point Attempt Percentage

| Coach | 2-Point Total | 2-Point % | Total Attempts |
|---|---|---|---|
| Jeff Ulbrich | 10 | 24.4% | 41 |
| Jonathan Gannon | 18 | 22.2% | 81 |
| Urban Meyer | 6 | 22.2% | 27 |
| Doug Pederson | 69 | 19.0% | 363 |
| Joe Judge | 9 | 18.8% | 48 |
| Chip Kelly | 2 | 2.7% | 73 |
| Marvin Lewis | 6 | 3.6% | 167 |
| Jon Gruden | 6 | 4.1% | 146 |
| Tom Coughlin | 2 | 4.5% | 44 |
| Kevin O'Connell | 7 | 4.8% | 145 |

All three are mixed effect logistic regression models using the coach of the team as a random intercept. We did some preliminary analysis into coach decision making and found that while there are overall trends we can see in when teams go for two-point conversions, some coaches are much more likely to go for two-point conversions than others (as seen in Table 2) and a random intercept allows us to account for this.

The first model uses time remaining in the game, the score differential as a continuous variable, and an indicator of if the team on offense is at home as predictors. We believe using the continuous score variable will allow us to identify bigger picture trends between score and going for two. The second model is the same as the first model, but instead uses a discrete score predictors which is a factor that considers some of the key scores we identified in Figure 2 (down 13, down 10, down 8, down 2, up 1, up 5) as levels compared to all other scores. This model will allow us to pick up on more specific trends related to how teams react to exact scenarios and how these change throughout the game. The last model is the same as model 2, but also includes the extra point percentage of the team's kicker to see if kicker strength changes coach decision making.

The next step of analysis is to create a model which tries to predict how likely a team is to succeed in converting a two point conversion. This is a logistic regression which predicts the outcome of a two-point conversion based on the play type (run vs. pass), time remaining, score, distance of the conversion (if there is a penalty, some conversions may be attempted from the one yard line and we would like to know whether or not teams should always choose to go for two in these situations), as well as an interaction between play type and distance (our hypothesis is that teams may be more likely to run the ball from the one yard line and defenses may adjust to this). Our model then uses the offensive team, defensive team, and quarterback as random intercepts to control for some of the variability between team and player strengths.

Lastly, we will use our model to try to predict how likely different team's are to succeed on two-point conversions in certain scenarios. Then, using the win probability that a team has when they are down 8 vs. 7 vs. 6 points, we will try to recreate a decision tree to answer how likely teams are to win if they use the go for two down 8 strategy that we laid out in the motivation. Using these win probabilities, our model, and the team's likelihood of converting extra points through the kicker's percentage, we can then aggregate to see if going for two in that scenario makes it more or less likely that the team wins.

## 4 Results

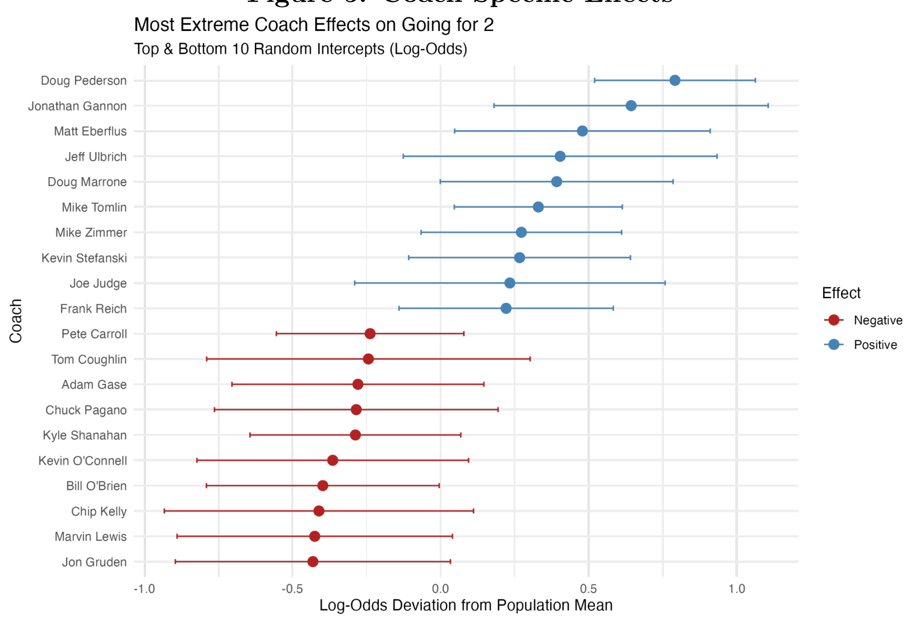
In our first logistic regression model, we predict how likely a team is to go for a two-point conversion using the time remaining in the game, score differential as a continuous variable, and home field advantage status with the team's coach (the decision maker) as a random intercept. As seen in Table 3, the score differential and time remaining in the game are both significant with negative coefficients. We are therefore able to identify the bigger picture trend of when teams go for two: as the score differential increases (meaning it becomes less negative for the team with the ball and they are either winning by more or losing by less), they are less likely to go for two) and as the time remaining in the game increases (as we get earlier in the game), teams are less likely to go for two.

Table 3: Model 1

| Variable | Estimate | Std. Error | CI Lower | CI Upper | p-value | |
|---|---|---|---|---|---|---|
| Intercept | -0.938 | 0.070 | -1.075 | -0.800 | 0.00 | *** |
| Time Remaining (sec) | -0.001 | 0.000 | -0.001 | -0.001 | 0.00 | *** |
| Score Difference | -0.049 | 0.003 | -0.054 | -0.043 | 0.00 | *** |
| posteam_typehome | 0.013 | 0.064 | -0.113 | 0.138 | 0.84 | |

We also can use this model to identify some of the specific coach effects of two-point decision making which is presented in Figure 5 below. As the chart shows, the effects can be fairly negative or positive. We identify a couple of interesting trends. First, the two coaches with the most positive effects are Doug Pederson and Jonathan Gannon, who both spent time as either a head coach or assistant coach with the Philadelphia Eagles, who are known for being very analytics heavy. Second, the coaches with the most negative effects like Jon Gruden, Marvin Lewis, and Chip Kelly are older-school, more traditional coaches. This may suggest something about how analytics vs. conservative and traditional styles of coaching impact decision making in these scenarios.

**Figure 5: Coach Specific Effects**




Our next model is the same model as Model 1, but we have changed score to a factor that specifically compares scores of down 13, down 10, down 8, down

2, up 1, and up 5 to all other scenarios. These are scores that we identified from Figure 2 as well as our own intuition. As the model shows, time remains a significant covariate, but with a much smaller coefficient than before. The score differentials we identified are almost all very positive and significant, with the most positive scenarios being down 2 (to tie the game), down 10 (to make it a 1 score game), and up 1 (to make it a field goal lead). We are somewhat surprised to find that a score differential of down 8 is close to 0 meaning it is almost identical to all other scenarios and in general coaches are not likely to employ the strategy of going for two down 8.

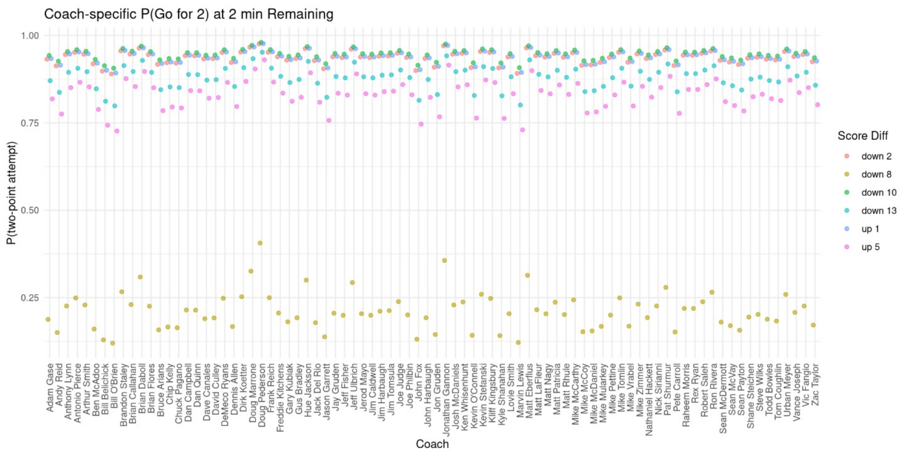
**Table 4: Model 2**

Fixed Effects: Estimates, 95% CI, and Significance

| Variable | Estimate | Std. Error | CI Lower | CI Upper | p-value | |
|---|---|---|---|---|---|---|
| Intercept | -1.309 | 0.083 | -1.473 | -1.146 | 0.000 | *** |
| Time Remaining (sec) | -0.001 | 0.000 | -0.001 | -0.001 | 0.000 | *** |
| Score Diff = –2 | 4.170 | 0.199 | 3.779 | 4.560 | 0.000 | *** |
| Score Diff = –8 | 0.087 | 0.181 | -0.268 | 0.443 | 0.630 | |
| Score Diff = –10 | 4.349 | 0.313 | 3.735 | 4.964 | 0.000 | *** |
| Score Diff = –13 | 3.457 | 0.351 | 2.769 | 4.145 | 0.000 | *** |
| Score Diff = +1 | 4.199 | 0.283 | 3.643 | 4.754 | 0.000 | *** |
| Score Diff = +5 | 3.058 | 0.163 | 2.738 | 3.377 | 0.000 | *** |
| Home Team | -0.146 | 0.072 | -0.286 | -0.006 | 0.041 | * |

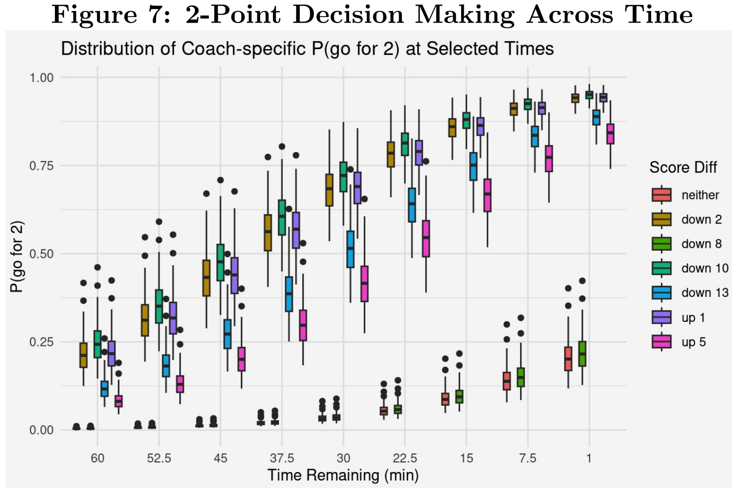
We also were interested in comparing how some of these trends vary across time and by coach. In Figure 6, we looked specifically at coaches' choices to go for two-point conversions in the closing minutes of a game, when model 1 tells us teams are more likely to go for two. As expected, coaches are much less likely to go for two points down 8 compared to all other scenarios, but we also can see that there is some variance across different coaches for the down 8 scenario (yellow dots). Some coaches are close to 0, while others like Jonathan Gannon and Doug Pederson are closer to a 0.5 probability. The up 5 scenario also has some variance, other scenarios like down 2 or up 1 are universally close to a probability of 1 in this scenario.

**Figure 6: 2-Point Decision Making by Coach and Score**




We next were interested to see how decision making in different score scenarios changes over time, which we present below in Figure 7. All scores tend to generally follow the trend from model 1, where they increase in probability over time. However, we can see that the uncertainty of decision making also changes across time and is highest in the middle of the game when coaches may be making different calls of how the game will play out. In late game scenarios, coaches tend to be more in line in taking two point conversions in these scenarios (although we can see that the uncertainty for the down 8 scenario does increase later in the game).

**Figure 7: 2-Point Decision Making Across Time**




We then went on to re-run this model with one extra predictor: the made extra point percentage of the team's kicker. We were curious to see if teams change their decision making based on reliability of their kicker. This is most interesting when a kicker gets injured in the game and the team has no available kicker, but unfortunately this is not something our model can account for with the data available to us. However, as Table 5 shows, we are still able to pick up on some of the impact of kicker strength. The coefficient, which is highly negative, is not necessarily super useful on its own given that the coefficient indicates the change in log odds when a kicker goes from 0 to 100 percent reliability which obviously is not the range we are looking at. However, it is highly significant and because the coefficient is so largely negative, the model does suggest that small increases in extra point percentage would have some decrease in a team's likelihood to try a two-point conversion which matches our intuition that they would be more likely to trust their reliable kicker.

**Table 5: Model 3**

Fixed Effects: Estimates, 95% CI, and Significance

| Variable | Estimate | Std. Error | CI Lower | CI Upper | p-value | |
|---|---|---|---|---|---|---|
| Intercept | 5.695 | 0.729 | 4.267 | 7.124 | 0.000 | *** |
| Time Remaining (sec) | -0.001 | 0.000 | -0.001 | -0.001 | 0.000 | *** |
| Score Diff = −2 | 4.155 | 0.201 | 3.760 | 4.549 | 0.000 | *** |
| Score Diff = −8 | 0.031 | 0.185 | -0.331 | 0.392 | 0.868 | |
| Score Diff = −10 | 4.316 | 0.318 | 3.694 | 4.939 | 0.000 | *** |
| Score Diff = −13 | 3.403 | 0.354 | 2.709 | 4.098 | 0.000 | *** |
| Score Diff = +1 | 4.247 | 0.286 | 3.686 | 4.807 | 0.000 | *** |
| Score Diff = +5 | 3.048 | 0.165 | 2.726 | 3.371 | 0.000 | *** |
| Home Team | -0.157 | 0.072 | -0.299 | -0.015 | 0.030 | * |
| XP Success Rate | -7.466 | 0.772 | -8.979 | -5.952 | 0.000 | *** |

After completing the analysis of when teams go for two, we attempted to model success rates on two-point conversions using score differential, time remaining, the play type (pass vs. run), the distance of the attempt (under certain unique penalty situation, the play can occur from the one yard line), an interaction between the play type and distance, the offensive team, and lastly a random intercept of the defensive team. We tried to use the team's quarterback as another covariate, but we did not have enough data points for each QB for this to work well. Table 6 shows the results of the model (excluding the team specific coefficients).

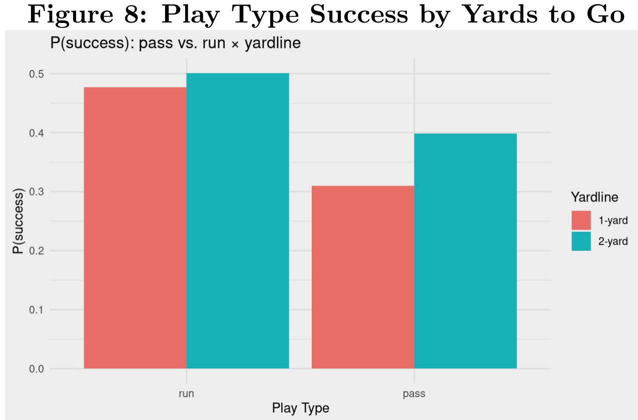
**Table 6: Model 4**

Fixed Effects: Estimates, 95% CI, and Significance (Team Effects Not Shown)

| Variable | Estimate | Std. Error | CI Lower | CI Upper | p-value | |
|---|---|---|---|---|---|---|
| Intercept | 0.035 | 0.332 | -0.616 | 0.686 | 0.917 | |
| Score Differential | -0.002 | 0.006 | -0.013 | 0.010 | 0.776 | |
| Time Remaining (sec) | 0.000 | 0.000 | 0.000 | 0.000 | 0.535 | |
| Pass Play | -0.417 | 0.151 | -0.714 | -0.120 | 0.006 | ** |
| One Yard to Go | -0.098 | 0.255 | -0.597 | 0.402 | 0.702 | |
| Pass × One Yard | -0.290 | 0.354 | -0.984 | 0.404 | 0.413 | |

There are a couple points of interest. First, we can see that run plays tend to be much more effective regardless of team. While teams may try to prevent runs in these close situations, it still may be more difficult to pass from so close as well. Next, we can see that the distance was not a strong predictor. This may partially be due to a lack of points from the one yard line, but we also find that it is because the overall success rate of conversion from the one yard line is actually slightly lower. In Figure 8 below, we also present how the play type success rate changes by yards to go. As we can see, especially on pass plays, it is much harder to convert from the one yard line (this intuitively makes sense because the windows are so tight). It is also slightly harder to run from the one yard line. This part of the model may need more data to give proper advice, but our initial takeaway is that it is not necessarily a good idea to change your strategy to go for a two just because there is a penalty.

**Figure 8: Play Type Success by Yards to Go**



The last piece of our analysis was to take a deep dive into the down 8 scenario using our model and win probability. We started by trying to estimate what the average win probability is for teams down 8, 7, and 6 in end of game scenarios. Unfortunately, it is challenging because win probability can change in different scenarios, and we want to get a broader picture, so we decided to filter to the last 10 minutes of a game and take the average win probability for teams, recognizing that it may sometimes be higher or lower. We ended up coming to estimates of 0.125 for down 6, 0.086 for down 7, and 0.072 for down 8. We then take these win probability estimates along with each team's predicted probability of converting a 2-point conversion in this scenario along with their extra point percentage rate to come up with an estimate of the expected win probability if the team goes for 2 or kicks an XP. Table 7 presents the teams with the highest and largest difference between E(WP) for 2-point converions vs XP. What we find is that while it varies based on a team's likelihood of converting, every team has a higher E(WP) if they go for two instead of kick an XP in this scenario. This is a very meaningful and interesting result given what our other model showed about the likelihood of going for 2 down 8, even in late game scenarios.

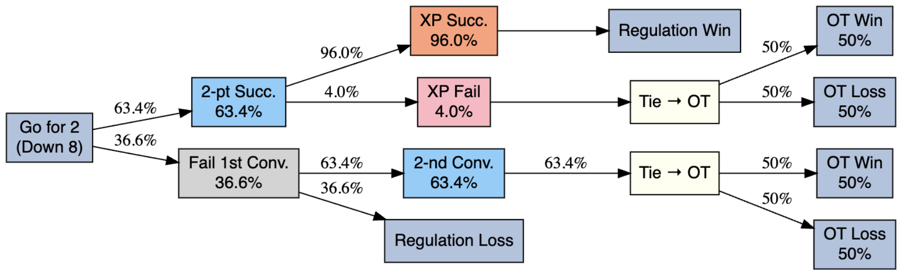
**Table 7: Team WP for XP vs. 2-Point Conversion**

Top and Bottom 5 Teams by Win Probability Gain from 2-Point Conversion (vs XP)

| Group | Team | WP (XP) | WP (2PT Run) | WP (2PT Pass) | ΔWP (Run - XP) | ΔWP (Pass - XP) | Max ΔWP |
|---|---|---|---|---|---|---|---|
| Top 5 | LAC | 0.0852 | 0.1099 | 0.1049 | 0.0246 | 0.0197 | 0.0246 |
| Top 5 | CLE | 0.0846 | 0.1068 | 0.1014 | 0.0221 | 0.0168 | 0.0221 |
| Top 5 | BUF | 0.0852 | 0.1072 | 0.1019 | 0.0220 | 0.0167 | 0.0220 |
| Top 5 | DAL | 0.0847 | 0.1062 | 0.1008 | 0.0215 | 0.0161 | 0.0215 |
| Top 5 | NYJ | 0.0851 | 0.1058 | 0.1004 | 0.0208 | 0.0153 | 0.0208 |
| Bottom 5 | WAS | 0.0849 | 0.0949 | 0.0897 | 0.0101 | 0.0048 | 0.0101 |
| Bottom 5 | ATL | 0.0854 | 0.0971 | 0.0917 | 0.0117 | 0.0063 | 0.0117 |
| Bottom 5 | TB | 0.0849 | 0.0970 | 0.0916 | 0.0120 | 0.0066 | 0.0120 |
| Bottom 5 | HOU | 0.0848 | 0.0970 | 0.0915 | 0.0122 | 0.0068 | 0.0122 |
| Bottom 5 | NO | 0.0855 | 0.0986 | 0.0931 | 0.0131 | 0.0076 | 0.0131 |

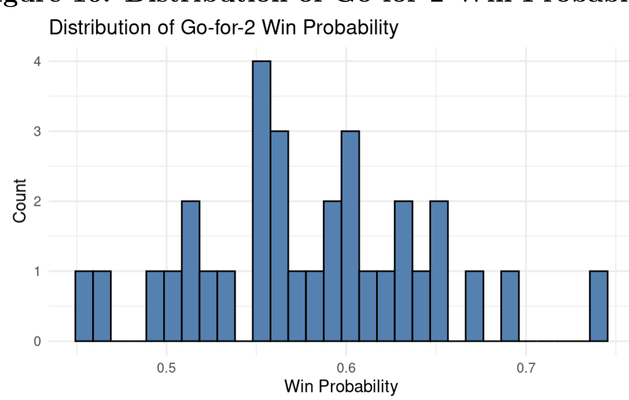
We also estimated this impact by assessing the probability of winning the game conditional on the team scoring another touchdown without the opponent scoring. This strategy is shown for the LA Chargers in the decision tree below. The strategy is to go for two after both touchdowns. For the LA Chargers, this gives a conditional probability of winning of approx. 74%. The going for 2 success estimates were calculated using our model, which estimates success rates for passes and runs, and the proportion of times that the teams have historically gone for passes and runs. This allows us to estimate the overall probability of making a 2 point attempt for each team. The extra kick success rate was estimated as a proportion of made attempts in the data.

**Figure 9: Decision Tree for the LA Chargers**



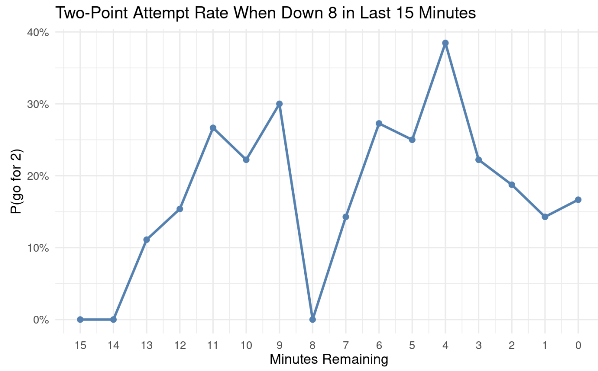

As a result, we found that only three teams had a conditional win probability of less than 0.5 by following this strategy. These teams were the TB Bucs (p=0.491), the Washington Commanders (p=0.461), and the Houston Texans (p=0.451). The highest conditional win probability was p=0.737 for the LA Chargers.

**Figure 10: Distribution of Go-for-2 Win Probability**




We saw in our initial EDA that the two point attempt rate when down 8 was less than 10%. We decided to investigate whether this changed towards the end of the game; we saw that this generally increases towards the end of the game, mainly fluctuating between 15 and 30 percent. There is not enough data to investigate these trends on a team-level, but the conditional probabilities suggests that coaches might be not taking enough 2-point attempts towards the end of the game when down 8.

**Figure 11: Two-Point Attempt Rate when Down 8 in 4th Quarter**




# 5 Conclusions and Discussion

Our results offer both some interesting conclusions and some areas for further investigation given our limitations. Firstly, we are able to do a good job of laying out the general trends in when teams are going for two-point conversions based on score, time, and specific scenarios. We also are able to provide evidence that kicker strength and coach specific thinking is a relevant factor in this process. We identify which coaches are have negative and positive effects in terms of how likely their team is to go for two. Some of the limitations of this first set of analyses are that we are not able to pickup some parts of within game variation (such as injuries to kickers) and we are also not able to account perfectly for how team strength on offense and defense impacts this decision (although we do control for teams by using coaches as a random intercept).

Next, we are able to lay out some predictions for how likely teams are to succeed. We believe that we provide a clear answer to the question of should teams go for it down 8, especially in late game situations, by showing through multiple analyses that teams increase their win probability by doing so. We also offer some evidence that two-point conversion after a penalty moves the ball to the one yard line are not necessarily any easier and therefore we would not recommend teams adjust strategy just because of a penalty. We also provide evidence for how play types may impact the likelihood of success.

There are some key limitations to our modeling of success. Our estimates that control for the team are looking at a 10 year period and don't account for teams changing within that period. The teams that our model predicts are most and least likely to succeed on two-point conversions could have had variations and ups and downs within that period. Additionally, we don't account for specific players on the field or how certain QBs, like Patrick Mahomes, may change this calculation and even how certain dominant defenders or injury ridden teams have different calculations. To improve our model, we could incorporate 4th down data from the 1 or 2 yard line and also try to isolate some differences across teams within the period of our data.