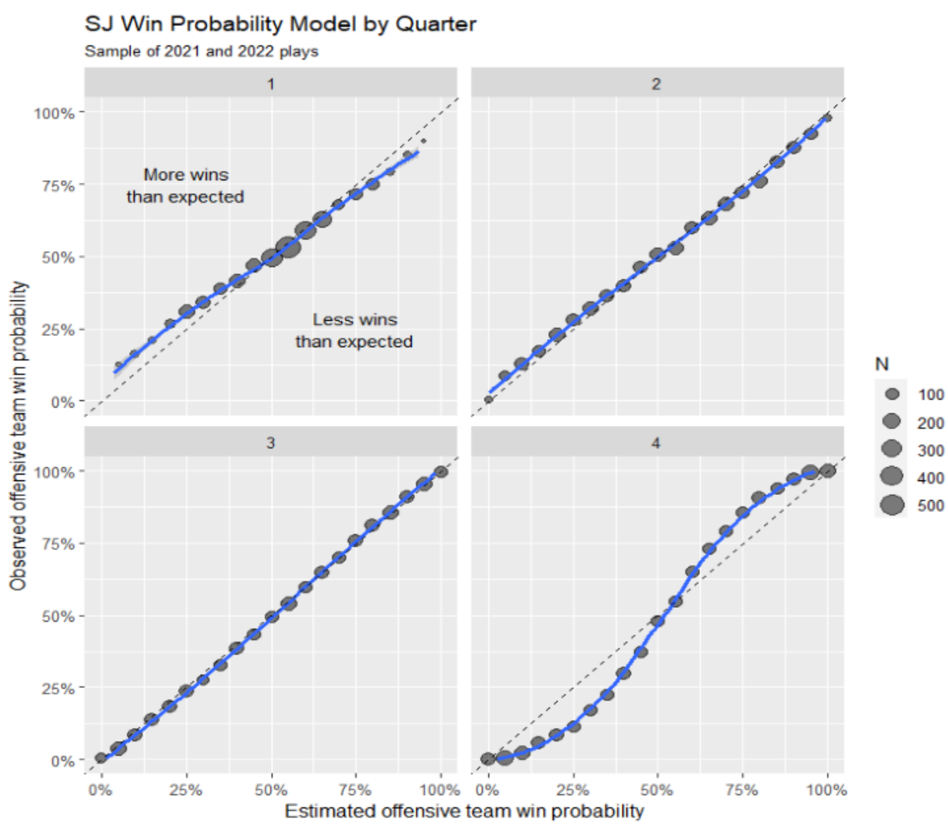
Scott Jue
MSDS 456
2/05/2023

Assignment 1: NFL Win Probability Model

1. Win Probability Model

To create the win probability model, I acquired NFL play-by-play data in R through the nflfastR library. Using the load_pbp() function, I obtained data for the 2021 and 2022 seasons, both regular and post season. I was reluctant to include the 2020 season data since that season was impacted by COVID-19 with players being in and out of the lineup. I then proceeded to add a winner field which had the winning team and a poswins which is a binary variable that has a “yes”/”no” value based on if the team with possession wins the game (Hill 2017). Next, I removed any observations with NA values in poswins and downs. These plays were related to kickoffs and extra point plays where there was no “down” or a “possession” team. In most cases those plays result in an immediate turnover, meaning that when the play ends, the other team gets possession. Therefore, these are not similar to typical football plays and I have chosen to remove them from the data. Additionally, I filtered the data to not included any non-plays such as plays that incur penalties and also any overtime plays since those rules are different than standard rules. Furthermore, I did not include the Super Bowl from the 2021 season, since that game is played at a neutral site I felt it was not similar to a typical regular or post season game. Once the data was cleaned and filtered, I was left with 77,450 plays for my analysis.

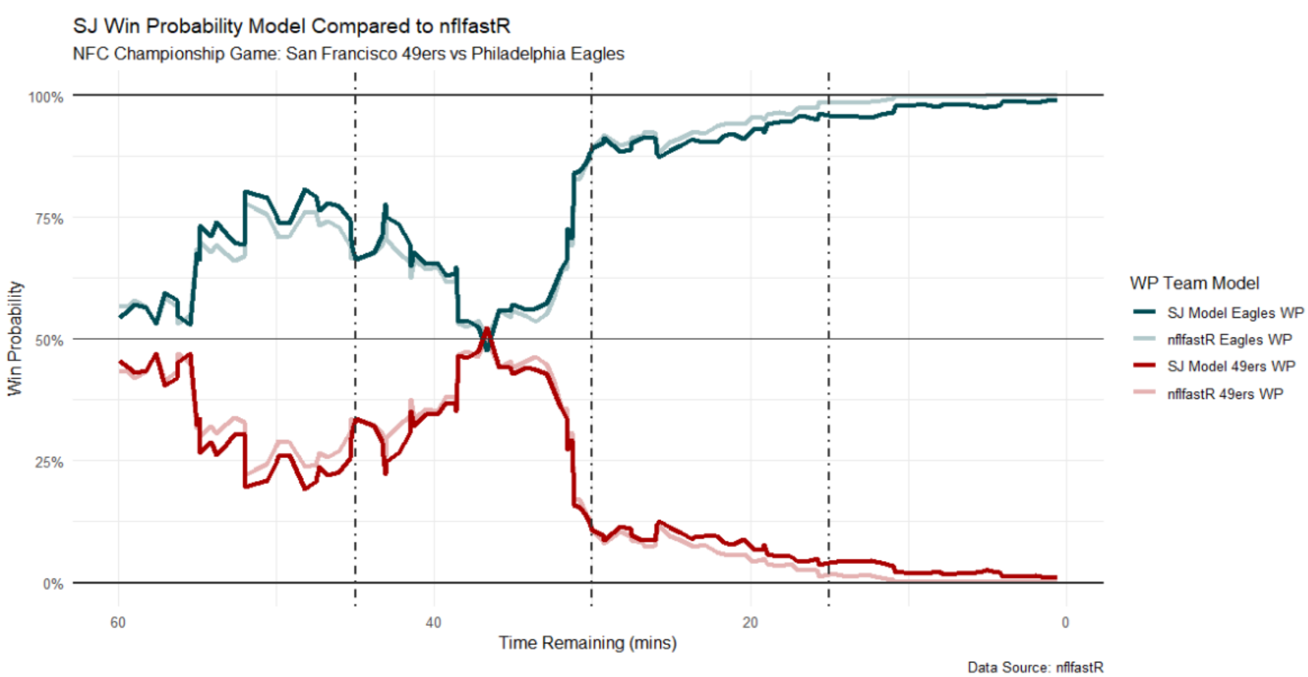
Using training data from an 80/20 train-test split, I used a generalized linear model approach using the glm() function in R to fit a model on the data. The variables selected for the model were quarter, down, yards to go for first down, time remaining in seconds, yards to go to opponents endzone, and score differential. These variables describe the time left in the game (quarter and game seconds remaining), the current score differential, spot of the ball on the field for the possession team (yards to go to opponents endzone), and current circumstances of play progression (down and yards to go). The approach and variables used for this model are similar to that of Hill (2017). Using the original win probability (wp) field from the nflfastR data as the actual value, I compared the accuracy of predicted values using RMSE. Using the unseen test data, the model produced a RMSE of 0.066 which I felt was very impressive. Because of how well the model performed, I did not choose to test different modeling approaches. Additionally, accuracy was also assessed using a similar approach used by Lopez (2017). Using the test data from the train-test split, I sampled 5 plays from each quarter in each game so that the games weighed equally. I then plotted these observations in below chart to show how well the model predicts these plays, aggregated by quarter. Then, to assess the accuracy, I looked at the projections in relation to the diagonal line. If there is a high degree of accuracy, then the projections will be closer to the diagonal line.



We can see that Quarters 2 and 3 are very accurate since pretty much all the points are on or touching the diagonal line. There is some variance in Quarter 1 the lower and upper ends of the win probabilities. Teams with low win probabilities win slightly more than expected and teams with higher win probabilities win slightly less than expected. However, the fourth quarter is where there is the most discrepancy in the model predictions, as there is a noticeable S-shape curve to the line. This indicates that the model over predicts teams with a low winning percentage and under predicts teams with a high winning percentage. Therefore, the model produces reasonable win probability results for Quarters 1, 2, and 3, but, will see a decrease in accuracy when predicting win probabilities in Quarter 4 particularly at the low and high win probabilities. If the model produces a win probability of less than .5 in the fourth quarter, then the actual win probability is likely lower. Alternatively, if the model produces a win probability of higher than .5 in the fourth quarter, then the actual win probability is likely higher.

2. Model Visualization of Individual Game

The game I chose to visualize using the win probability model was the NFC Championship game between the San Francisco 49ers and Philadelphia Eagles. When training the model, I specifically held this game out of the training set so that there were no overlapping plays in the test and training data to avoid overfitting. As such, I could have a reasonable comparison of how the model predicted win probability compared to the nflfastR computed win probability.



We can see that the model tracks (darker shaded colors) very closely to nflfastR's win probability (lighter shaded colors) for both the 49ers and the Eagles. Some variance is observed in the first 20 minutes of the game where the Eagles win probability was predicted higher than nflfastR's win probability and the 49ers win probability was predicted lower than nflfastR's win probability. Additionally, there is some variance in the later parts of the game where the model under predicted the Eagles' win probability and over predicted the 49ers' win probability. Next, I will look at the plays that had a significant impact on this game's win probability model.

The Eagles had a higher win probability of .54 to begin the game as they were the number 1 seed in the NFC and also playing at home. In their first possession, they went for it on 4th and 3 and completed 29 yards pass to put them at the 6yd line. This increased their win percentage by .15, from .53 to .68. The eventually scored a touchdown a few plays later which increased their win probability to .73. On the next drive, when the 49ers had possession, they 49ers turned it over on a fumble by Brock Purdy who subsequently got injured on this play. This gave the Eagles the ball back around mid-field and increased their win probability to .8. The teams punted back and forth to each other for their next couple of possessions, which effectively made both team's win probability converge towards 0.5. This is because of the time remaining in the variable. As time went by and the score differential remained at 7, the Eagles win probability decreased slightly since they weren't adding to their lead. The 49ers win probability slowly increased since they were able to keep the game close during this time keeping it to a "one possession". About halfway through the second quarter, the 49ers scored a touchdown to even things up which decreased the Eagles win probability to .53 and the 49ers increased to .47. On the next drive the Eagles were faced with a 4th and 1 situation. Surprisingly, to start this play the 49ers had the higher predicted win probability of .52 and the Eagles had .48. However, after going for it again on fourth down and getting a first down, the Eagles win probability increased to 0.56. They eventually scored a touchdown on this drive further increasing their win probability to .73. On the next drive the

49ers fumbled the ball after 2 plays to give the Eagles the ball back at the 49ers 30-yard line, increasing the Eagles' win probability to .84. Scoring a touchdown a few plays later and adding to the lead to make it 21-7 increased this probability to .88. Mid-way through the third quarter the Eagles scored another touchdown which increased their win probability to .95. On the next drive the 49ers went for it on 4^{th} down and did not convert, so they turned it over on downs at mid-field which increased the Eagles win probability to .98. The Eagles then scored a field goal with 5:17 remaining in the game to make it 31-7 but this scoring play did not materially change the win probability for either team as this play did not really have any significant influence on the outcome of the game. The game was essentially already decided at this point.

One question that arises when looking at this game is does the model account for all plays that have a significant impact on the outcome for the game? In my opinion, there were two plays that significantly influenced this game that I believe the model did not fully capture. The first happened in the first quarter during the first turnover by the 49ers where Brock Purdy injured his elbow which led him to fumbling the ball. This injury knocked him out of the game and the 49ers were forced to bring in their 4^{th} string QB, Josh Johnson. The second play was similar in nature in that Josh Johnson suffered a concussion at the beginning of the third quarter which knocked him out of the game. Brock Purdy was forced to come back in, but due to his injury he was not able to throw far and the 49ers essentially were reduced to only running plays and short screen passes. In reality, those plays and events should increase the win probability of the Eagles, since the 49ers were essentially playing without a quarterback for almost half the game.

Modeling the NFC Championship game gives us insights into what kind of plays significantly impact win probabilities. The conversion of the 4^{th} and 3 attempt by the Eagles that increased their win probability by .15 seemed to be a play that highly influenced the game because it set them up to score a touchdown on the opening drive and gave the Eagles the early momentum. This would make it appear that converting on a 4th down attempt could provide huge swings in the game. However, this play in particular is a bit misleading due to it being a 29-yard completion at the 49ers 6-yard line. In other circumstances, a 4th down conversion might not have such a huge effect. For example, a team that goes for it on a 4th and 1 at mid-field and chooses a running play to run up the middle for a 2-yard gain. This play doesn't have as much influence on the win probability as the play from the NFC Championship game. Additionally, the model's predicted winning probability for the Eagles didn't increase materially when they scored a field goal. This would make it appear that field goals do not have an impact on the game. However, this is misleading because due to the time remaining and score differential, this scoring play did not have any impact on the outcome of the game. If the score differential was smaller or if the teams were tied, then a field goal would significantly impact the win probabilities.

This model produced a few learnings that can be helpful to a football team. First, it shows how turnovers can impact win probabilities, especially turning it over in your own defensive zone. More importantly, it shows that scoring off of a turnover can provide significant increases in win probabilities. Secondly, it showed how 4th down conversion plays can produce significant increases in win probability depending on the location on the field. Lastly, it shows how getting an early lead by halftime, even if its just 2 touchdowns, can give the home team over 80% win probability. Therefore, the home team might be more inclined to be aggressive with their play calling earlier in the game instead of later in the game.

3. Limitations and Future Directions

Though this model produces reasonably accurate results, there are several limitations to this model. First, not all plays are included in this model, so it is not able to predict win probability impact from plays that incur a penalty, kickoff plays, extra point plays or 2-pt conversion plays. Although, the outcomes of extra point and 2-pt conversion plays are captured in the score differential in the next play. Additionally, Superbowl data was removed from the model training data set as these are unique games, so this model may not perform as well when predicting win probabilities of Super Bowl games. Another limitation of this model is that overtime plays were not included in the model training dataset. Overtime has a different set of rules and can also turn into a sudden death format. For this reason, this model should not be used to predict win probabilities for any plays that occur during overtime. As mentioned in the discussion about the NFC Championship game between the Eagles and 49ers, the model also does not account for injuries which could significantly influence the outcome of the game depending on the position and severity of the injury. In that case of the 49ers, having your 3rd string and 4th string quarterback get injured in the middle of the game severely changes what plays you can run and the team's offensive potential. Similarly, the model does not account on-field personnel, play calls, schemes, or formations all of which also have in impact on the game and team performance. For example, in the previous example, when both 49ers quarterbacks were injured, the majority of their playbook was reduced to running plays. The likelihood of overcoming a 21-point deficit by only running is pretty slim. To overcome it when you opponent knows you can only run and therefore can play run defense, is even slimmer. So, I would consider including these variables to further improve the model in the future. Furthermore, I would perform further analysis on quarter plays and their impact on win probabilities since the model is not as accurate in that fourth quarter. Perhaps plays in the fourth quarter could be weighted differently than the first three quarters.

This model is comparable to the nflfastR win probability model with the exception to the fourth quarter win probability predictions where the model over predicts teams with lower winning probability and under predicts teams with higher winning probability. Another publicly available model is provided by numberfire.com (Appendix A). This model takes in to account team strength, score, time left, and possession. When comparing the two models using the NFC Championship game example from above, numberFire's model is more reasonable due to the team strength variable. To start the fourth quarter numberFire's model gives the 49ers a 1.24% win probability, whereas my model gives them a 4.38% win probability. Although both models produce significantly low probabilities, the team strength variable likely accounts for the injuries suffered earlier in the game and therefore, adjusts the win probability lower than my model. Lastly, different modeling approaches have been used such as a random forest model by Lock and Nettleton (2014) that use down, score, game seconds remaining, adjusted score, Las Vegas pre-game point spread, time outs remaining for the offense, time outs reaming for the defense, total points scored, yards from own goal line, and yards to go for first down. The accuracy of this model has been evaluated by Lopez (2017) have shown to produce fairly reasonable results (Appendix B). Therefore, for future work I would consider different modeling approaches such as random forest models and adding additional predictor variables such as team strength and Las Vegas pre-game point spread.

Possible other applications of this model include using the predicted win probability to determine if a team should go for it on fourth down or if a scoring play has little impact to the outcome of the game. For example, when the Eagles went for it on 4th down and converted, their win predicted win probability increased by 0.15. This win probability also got a boost from the play resulting in a 29-yard catch to put the Eagles in the red zone. However, in situations like this we can potentially see if going for it on 4th down results in a significant change in win probability and determine if it is worth or not. If the team doesn't convert, we can simulate a turnover on downs and see what the win probabilities look like if the opposing team gets the ball at that spot. Additionally, the Eagles' late field goal didn't do much in terms of change win probability percentage, so perhaps there are situations where logically it makes sense to score more points, but in reality, it may not affect the outcome or win probability. Since the field goal did not materially improve the Eagles' win probability, maybe there is an alternative play that was less risky than kicking, such as punting or just running it and possibly turning it over on downs knowing the 49ers had trouble moving the ball. Maybe the kicker hasn't been performing well recently and there is a low chance of making the field goal or potentially having it blocked and returned for a touchdown. So, knowing that a field goal does not increase the win probability, the team may opt to not kick the field goal; therefore, models like this one could facilitate in-game play calling to maximize the team's win probability.

4. Additional Relevant Resource Topic and Potential Implementation Plan

As mentioned above, one limitation of the win probability model is that it does not account for extra point or 2-point conversion plays since there is no down associated with these plays. Chapter 23 of *Mathletics* discusses how to build and implement a chart that summarizes a team's optimal extra-point/2-point conversion attempt strategy based on the score and time remaining left in the game. Two approaches were discussed to develop this chart. The first approach assumes there are 10 possessions or less in the game to build a probability model using the probability of each possible outcome on the game's last possession (field goal, touchdown, or no score) multiplied by the probability of winning given each possible outcome of the possession. The second approach developed by Morris (2017) from FiveThirtyEight is a bit more simplified in that it builds off of a previous win probability model and uses incremental increases in points to determine changes in win probabilities. An example of this chart can be seen in Appendix C. This chart assumes there are 10 minutes or less in the game and displays the change in win probability based on what impact the extra-point kick or the 2-point conversion has on the lead change. Additionally, this chart includes a description of what the lead change means in terms of tactical benefit. Therefore, this would be the approach I would recommend to implement alongside the win probability model.

Adopting this type of chart would not require any additional data since it builds off the current win probability model. An analytic tool that uses the methodology of from this chart could be used in game to recommend if the team should kick the extra point or go for 2. Building this type of tool would require the analytics department to add this form of analysis to the current win probability model and to develop an algorithm that recommends going for 1 or 2 based on current game data such as time remaining to ensure we are under 10 minutes remaining in the game, as well as, the current score as inputs. This recommendation can be communicated to the coaching staff or this information can also be displayed in a dashboard on a tablet. The dashboard would display the recommendation and could also provide specific details such as the change in win probability or even the full chart. The team could then use this analytical tool to decide if it is more favorable for the team to kick an extra-point or go for a 2-point attempt upon scoring a touchdown with less than 10 minutes remaining.

References

Hill, Stephen. 2017. “Building a Basic, in-Game Win Probability Model for the NFL.” *Medium*. December 29. https://medium.com/@technocat79/building-a-basic-in-game-win-probability-model-for-the-nfl-54600e57fe1c.

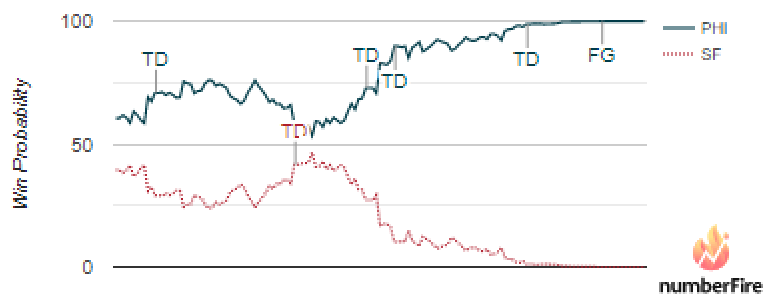
Lock, Dennis, and Dan Nettleton. 2014. “Using Random Forests to Estimate Win Probability before Each Play of an NFL Game.” *Journal of Quantitative Analysis in Sports* 10 (2). doi:10.1515/jqas-2013-0100.

Lopez, Michael. 2017. “All Win Probability Models Are Wrong - Some Are Useful.” *StatsbyLopez*. March 11. https://statsbylopez.com/2017/03/08/all-win-probability-models-are-wrong-some-are-useful/.

Morris, Benjamin. 2017. “When to Go for 2, for Real.” *FiveThirtyEight*. February 3. https://fivethirtyeight.com/features/when-to-go-for-2-for-real/.

<u>Appendix A: numberFire Win Probability Model of NFC Championship Game (Eagle vs 49ers)</u>

100.0% WIN PROBABILITY



End Game

HOW IS THIS CALCULATED?

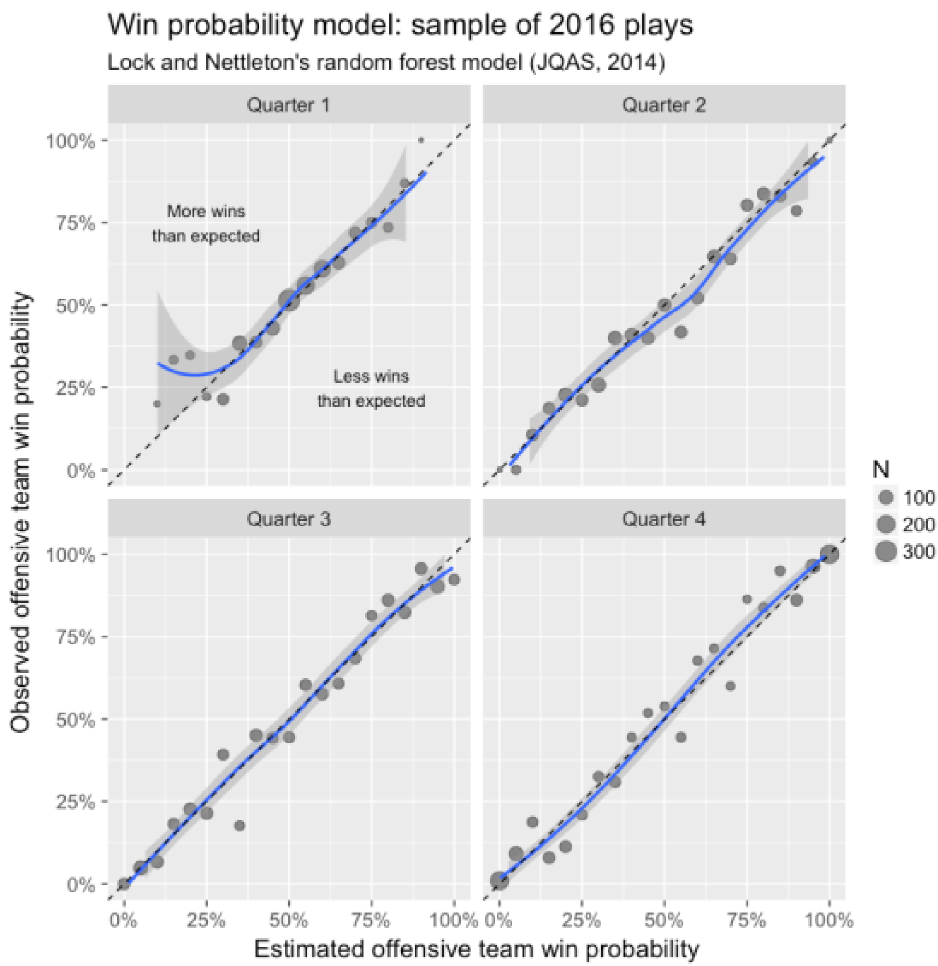
WIN PROBABILITY (WP)

Represents each team's chance to win the game.

Takes Into Account

Team strength, score, time left, possession.

Appendix B: Locket & Nettleton Win Probability Model



Appendix C: Morris Lead Change Win Probability Impact Chart

GOING FROM A LEAD OF	... GIVES YOU THE TACTICAL BENEFIT OF	... AND A WIN PROBABILITY CHANGE OF
0 to 1	The lead	+8.4
1 to 2	A generic point	+1.8
2 to 3	Puts you up a field goal	+6.5
3 to 4	Puts you up more than a field goal	+5.0
4 to 5	A generic point	+2.9
5 to 6	Puts you up two field goals	+3.1
6 to 7	Puts you up a touchdown	+5.2
7 to 8	Puts you up a TD with a 2-point conversion	+3.3
8 to 9	Puts you up two scores	+2.9
9 to 10	Puts you up by a touchdown and a field goal	+2.2
10 to 11	Puts you up by a TD with a 2-point conversion and a FG	+1.3
11 to 12	Puts you up more than a touchdown and a field goal	+1.1
12 to 13	Puts you up a touchdown and two field goals	+0.4
13 to 14	Puts you up two touchdowns	+1.0
14 to 15	Puts you up two TDs with a 2-point conversion	+0.5
15 to 16	Puts you up two TDs with two 2-point conversions	+0.7
16 to 17	Puts you up three scores	+0.2

Win probability changes apply to the listed situations when there are 10 minutes remaining in the game.