Evaluating the Impact of Special Teams on Winning

Matthew Ladden — Trinity School '25
&
Ethan Seung — Harvard-Westlake School '25

January 28, 2024

Abstract

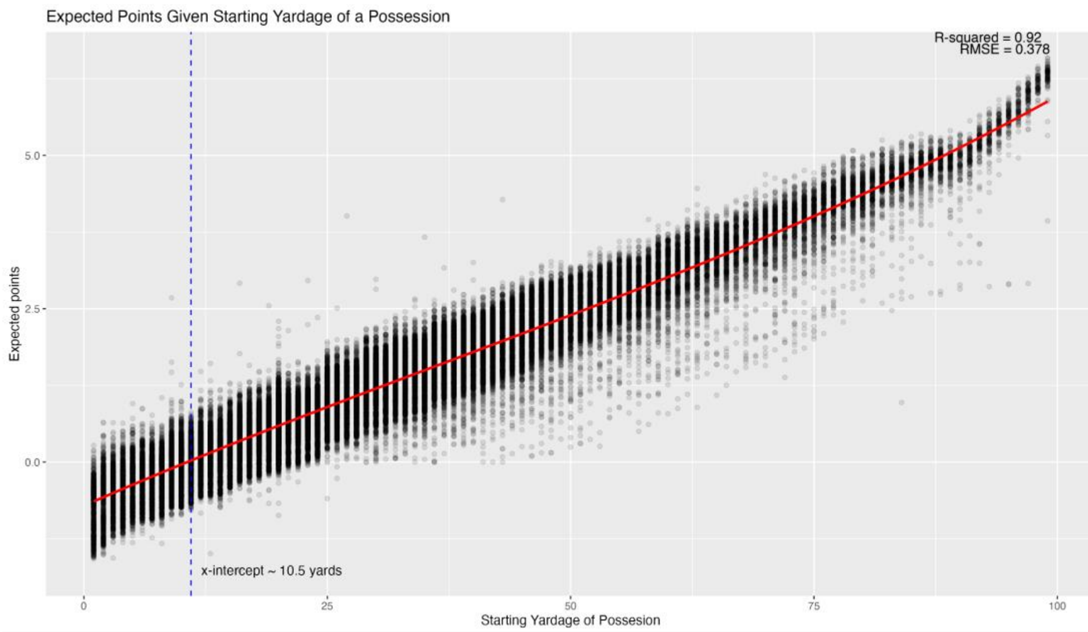
This paper evaluates the per-play impact of special teams on the winning percentage of NFL teams compared to offense and defense. Employing data from the 1999-2022 NFL seasons in the nflfastR database, we developed a new calculation for Expected Points Added (EPA) to accurately assess the contribution of special teams plays, including punts, kickoffs, and field goals, independent of coaching decisions. This methodology allowed us to create our Special Teams Performance Index (STPI), a composite metric that consolidates the performance of all special teams units on a team. We then used STPI in combination with offensive and defensive ratings from ProFootballReference in a regression model to evaluate the importance of special teams on winning percentage. Finally, since there are fewer plays on special teams than on offense or defense in a game, we adjusted the resulting numbers to find the impact on a per-play basis. We found that an individual play on offense is 3.6 times and that an individual play on defense is 2.6 times more impactful on winning percentage than a play on special teams. Our results can be extrapolated to coaching decisions concerning playing starters on special teams. A special teams play has a significantly lower impact on a game as compared to one on offense or defense, and plays on special teams have a significantly higher per-play injury rate; therefore, starters on offense and defense should generally not be used on special teams.

Introduction:

Special teams plays, including punts, kickoffs, field goals, and extra points, can significantly influence a football game's outcome. While extra points are limited almost exclusively to a make or a miss, all other special teams plays can affect the score of a game through the resulting field position. For example, a missed field goal can set up the opposing team with a field position that makes scoring points easier. Similarly, pinning a team back at their two-yard line after a punt undoubtedly makes it harder for them to score, but just how impactful is this play in the course of a game? George Allen, a Pro Football Hall of Fame coach, once asserted that: “Football is one-third offense, one-third defense, and one-third special teams,” but his commentary was not grounded in analytics. Research has been published concerning the importance of punting by the 33rd Team, but, according to this model, the top five punting units all had negative expected points added (EPA). Like many others, this EPA model took into account coaching decisions of whether to punt or go for it when calculating the EPA for a given play. Such decisions, when statistically incorrect (i.e. the coach should have elected to go for it), automatically lower the EPA of a punt, regardless of the outcome of the play. Our research aimed to determine the effect that the entirety of a special teams unit can have on a team’s winning percentage relative to its offensive and defensive units on a per-play basis, independent of erroneous coaching decisions.

Materials and Methods:

An EPA model was used to isolate and measure the performance of special teams units, which measures the actual point value of a given play compared to the expected points derived from historical averages; however, as described earlier, the EPA models that come in datasets like nflfastR, account for incorrect coaching decisions like punting/attempting field goals instead of going for it, rather than directly measuring on-field player performance. Thus, we calculated a new form of EPA by creating an expected points (EP) model to determine the value of a special teams play solely based on the starting field position, eliminating the impact of coaching decisions. This was accomplished by plotting the starting field position of every drive in the dataset along with each drive's result (in points). Our EP model would serve as the baseline for calculating EPA. Below is our cubic regression model of EP vs. Starting field position from the 1999-2022 seasons, which consisted of 79,389 total possessions:



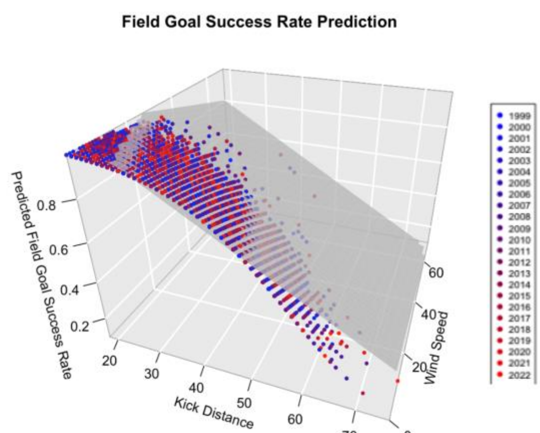
Kicking it off with kickoffs. We analyzed kickoffs using data from the 2016-2022 NFL seasons, a period chosen because of a rule change that moved the touchback from the 20-yard line to the 25-yard line. To calculate the EPA for kickoffs, we first determined the average resulting field position after a kickoff, which was the 25.52-yard line. Using a version of our EP model with only data from the 2016-2022 seasons, we calculated the expected points a team would score from this average field position. We then calculated the expected points value from the actual starting field position of each kickoff. Taking the difference of these two values gave us the EPA for each kickoff for the kicking team. The EPA for the receiving team is simply the negative of this value. We then calculated the average EPA per kickoff for both kicking and receiving teams for each season and analyzed how these averages correlated to the teams' winning percentages.

To calculate the EPA for each punt, we used the average resulting field position for each yardage where a punt was kicked from. Due to small sample sizes at some of the yardages, we calculated this by averaging the net positions of punts kicked within two yards of each other. For example, to determine the average resulting position for a punt from the 3-yard line, we used data from punts kicked from the 1-yard line to the 5-yard line. We applied our EP model to each average result to estimate expected points for a punt at every yardage. We then calculated the expected points for each punt's actual resulting field position. The difference between these two expected points values gave us the EPA for the punting team for each punt. By taking the negative value of this EPA, we obtained the EPA for the punt return team. We averaged the EPA of all punts for each team and each season and then analyzed how these averages correlated with the teams' winning percentages.

To calculate the EPA of a field goal attempt, two factors must be evaluated: the expected points for the field goal and the impact of the field position given to the opponent on a miss. The expected points for a field goal is the likelihood of success multiplied by three, making the EPA for any successful FG:

$$EPA_{Made\ FG} = 3 - (3 * Probability_{field\ goal\ make})$$

To calculate the probability of a successful field goal, we created a regression plane, factoring in kick distance and wind speed for field goal attempts from 1999-2022:



On a missed field goal, the value of the field position given to the opposing team is calculated by comparing the field position's expected points value to the expected point value of the average starting field position after a kickoff (the result if the field goal had been made). Thus, the EPA for a missed field goal is the negative value of the expected points for the field goal minus the difference between the expected points of the resulting field position and the average starting position from kickoffs:

$$EPA_{Missed\ FG} = -(3 * Probability_{field\ goal\ make}) - (EP_{field\ position} - EP_{average\ field\ position\ from\ kickoffs})$$

We did not factor in the expected points of the field position given to a team after a field goal miss if all possessions in the same half following the missed field goal have a combined likelihood of less than 5% chance of scoring (as per nflfastR's no_score_prob), such as when there is insufficient time left for the opposing team to score or when a team chooses to kneel. We then calculated the average EPA per field goal for each team in every season between 1999-2022 and analyzed how these averages correlated to the teams' winning percentages.

To compare the impact of special teams units as a whole, we created our own Special Teams Performance Index (STPI). This was created using the sum of the standardized average EPAs for each special teams unit on every team in every season. Using the standardized Offensive and Defensive Simple Rating Systems (OSRS and DSRS) from ProFootballReference and a standardized STPI, we ran a multiple logistic regression model to find the impact of offense, defense, and special teams on winning percentage. We also ran a regression model without the inclusion of our STPI to compare the difference in root mean square error (RMSE) and see if there was an improvement in our rating. By using STPI and thereby including special teams plays that only encompass 15% of all plays in NFL games, we improved the RMSE as compared to the model with only OSRS and DSRS by 1.9%, from 0.107 to 0.105. Taking the ratio of the coefficients from the logistic regression model would allow us to compare the impact of special teams to offensive and defensive units separately as it pertains to winning; however, since special teams units are on the field less than the offense and defense, we divided our results by the ratio of plays each unit is involved in (excluding QB kneels and spikes) to find the difference in per-play impact on winning percentage.

Results and Conclusion:

Results from Multi-Regression Model:

Coefficients:	Estimate:
z_OSRS	0.55645
z_DSRS	0.40596
z_STPI	0.02877

Number of Occurrences per Play Type:

Play Type:	Number of Occurrences:	Percentage:
Pass	463,377	48.68%
Run	340,589	35.78%
Kickoff	63,860	6.71%
Punt	59,644	6.27%
Field Goal	24,356	2.56%
Total	951,826	

The top left table contains the results from our multi-regression model, with each coefficient representing their impact on winning percentage. From this, we determined that offense and defense were 19.3 times and 14.1 times as impactful over the course of a game as special teams, respectively. To compare the impact on a per-play basis, we divided the results from the regression by the ratio of plays on special teams to offense/defense, given by the top right table. We found that one play on offense and defense is 3.6 times and 2.6 times more impactful on a higher winning percentage than a special teams play, respectively.

Our results can be extrapolated to coaching decisions on deciding if starters should play on special teams. Special teams has a significantly reduced per-play impact on a game compared to offense and defense and a significantly higher per-play injury rate; therefore, starters on offense and defense should generally not be used on special teams.

A further extension of this analysis would be to isolate the kickoff or punt unit from the return unit on each play, using advanced tracking data. This would allow for assessing how much of the resulting EPA stems from the performance of the kicking teams versus the return team, rather than the combination of both units on the field. Finally, the new kickoff rules passed by the NFL have the potential to alter how kickoffs are executed and valued. By moving touchbacks up to the 30-yard line, the NFL will encourage more returns on kickoffs, increasing variability in results and the importance of the play to both the return and kicking units. It will be interesting to rerun this analysis next year on the 2024-2025 NFL season data to determine how this change affects the value of kickoff units and special teams as a whole.

Acknowledgments:

We would like to thank our partners from Wharton's Moneyball Academy, Jonathan Dobkin (Mamaroneck High School '24), Tyler Frick (Palo Alto High School '24), and Xinhang Chen (Renert School '24), for their invaluable contributions in conceptualizing, researching, and analyzing the data that formed the basis for this paper. We would also like to extend our gratitude to Professor Wyner, Ryan Brill, Dylan Perlstein, Zeke Kelz, and Elan Roth for their help and instruction throughout the Wharton Moneyball Academy program.

References:

Ainbinder, R. (2020, September 11). The Future of Special Teams in the NFL. Sports, Tech, Biz. https://www.sportstechbiz.com/p/the-future-of-special-teams-in-the

Horowitz, M., Yurko, R., & Ventura, S. (n.d.). An R package to quickly obtain clean and tidy NFL play by Play Data. Dev status. https://www.nflfastr.com/

NFL team points per game. NFL Football Stats - NFL Team Points per Game | TeamRankings.com. (n.d.). https://www.teamrankings.com/nfl/stat/points-per-game

Nflverse. (n.d.). Nflverse/nflverse-DATA: Automated nflverse Data Repository. GitHub. https://github.com/nflverse/nflverse-data

Simpson, R., & Engler, P. (2023, November 14). Study: Impact of special teams on NFL Success. The 33rd Team. https://www.the33rdteam.com/case-study-impact-of-special-teams-on-nfl-success/

Tokarz, D. (2019, February 18). Beyond Make Percentage: Evaluating NFL Kickers. Yale Undergraduate Sports Analytics Group. https://sports.sites.yale.edu/nfl-kicker-evaluation

Pro Football Stats, history, scores, standings, playoffs, Schedule & Records. https://pro-football-reference.com/