The Value of Punting Units in the NFL

Logan Donaldson, Justin Nam, Sam Oberly, Gaby Quintana

Johns Hopkins University | Whiting School of Engineering | Baltimore, MD

Design Day 2021

Introduction

In collaboration with the Baltimore Ravens our group was tasked with determining the value of punting units in professional football. In the past decade NFL teams have readily embraced analytics to inform their game-day decisions. High leverage offensive and defensive situations such as deciding to go for it on fourth down have been subjected to immense numerical scrutiny. However, outside of placekickers, special teams remains an analytical blind spot. To remedy this discrepancy, we conducted exploratory data analysis on a play-by-play NFL data set (some data courtesy of PFF) spanning the years 2017-2019 with a focus on punts and punt returns.

Objectives

- Can the length of a punt return be accurately predicted from in game variables?
- Does a long punt return create quantifiable momentum on the ensuing drive?
- Are some punt or return units significantly better than their peers?

Results

```
Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              -1.48848    1.83077  -0.813 0.416263
clean_catch_binary        5.44655    0.51914  10.491  < 2e-16 ***
actual_kick_direction_R   2.48641    0.58347   4.261 2.09e-05 ***
actual_kick_direction_C   2.68449    0.46713   5.747 1.00e-08 ***
return_direction_R        3.21779    0.48157   6.682 2.81e-11 ***
return_direction_C       -0.27768    0.42263  -0.657 0.511219
hang_time                -3.43597    0.42885  -8.012 1.60e-15 ***
kick_yards                0.31508    0.02654  11.872  < 2e-16 ***
vise_count                0.80821    0.22800   3.545 0.000399 ***
penalty_punt             -6.07556    0.48293 -12.581  < 2e-16 ***
penalty_return           -1.35029    0.78250  -1.726 0.084521 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.376 on 2971 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.1517, Adjusted R-squared:  0.1489
F-statistic: 53.14 on 10 and 2971 DF,  p-value: < 2.2e-16
```

Figure 1—Linear Regression Model with Return Yards as Response Variable
Starting with 13 quantitative and 7 categorical predictor variables the model was reduced to the above. The near zero p-value suggests that the predictor variables and response variable are indeed correlated. The model suggests that a clean catch, direction of punt, direction of return, punt distance, punt hang time, whether a penalty occurred on the play, and the number of vise who are sent downfield to tackle the returner, all have a significant effect on the distance of a punt return. However, the adjusted R-squared value of 0.1489, indicates that much of the variance in the return yards remains unexplained by the predictors.

Results

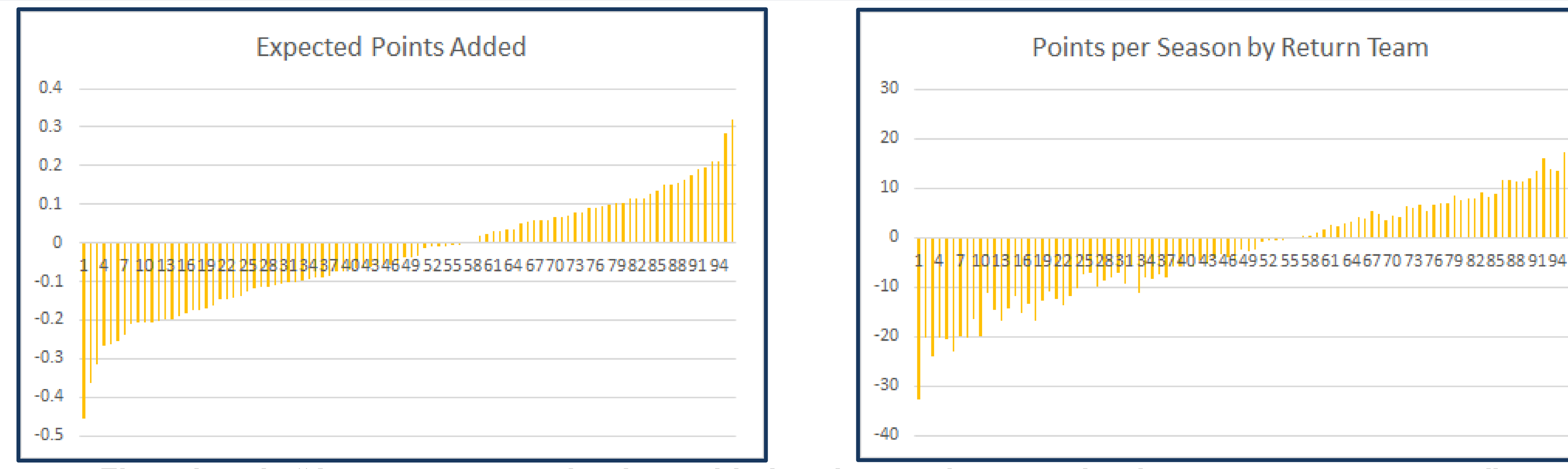


Figure 2—First chart is "Average expected points added to the punting team by the return team per punt" Second chart is "expected points added to return team over course of a season"
A negative value means that the return team had a negative impact on their opponent, or equivalently, a positive impact on their team. Over the course of a season, an average return team can expect to decrease their opponents expected points by a 2.825 points. This is rather insignificant over an entire season- for example, the ravens let up 303 points in 2020, so a >3 point difference over the course of a season is not much. The best return teams (top 8 of 96 between 2017 and 2019) can cost the opponent upwards of 20 expected points per season (which is almost 7% of the points the Ravens gave up). The worst teams (11 of 96) can help the opponent by upwards of 11 expected points over the course of the season, which is small but may be noteworthy. A top punt return team can noticeably decrease the opponents expected points, whereas an average return team makes very little difference over the course of a season.

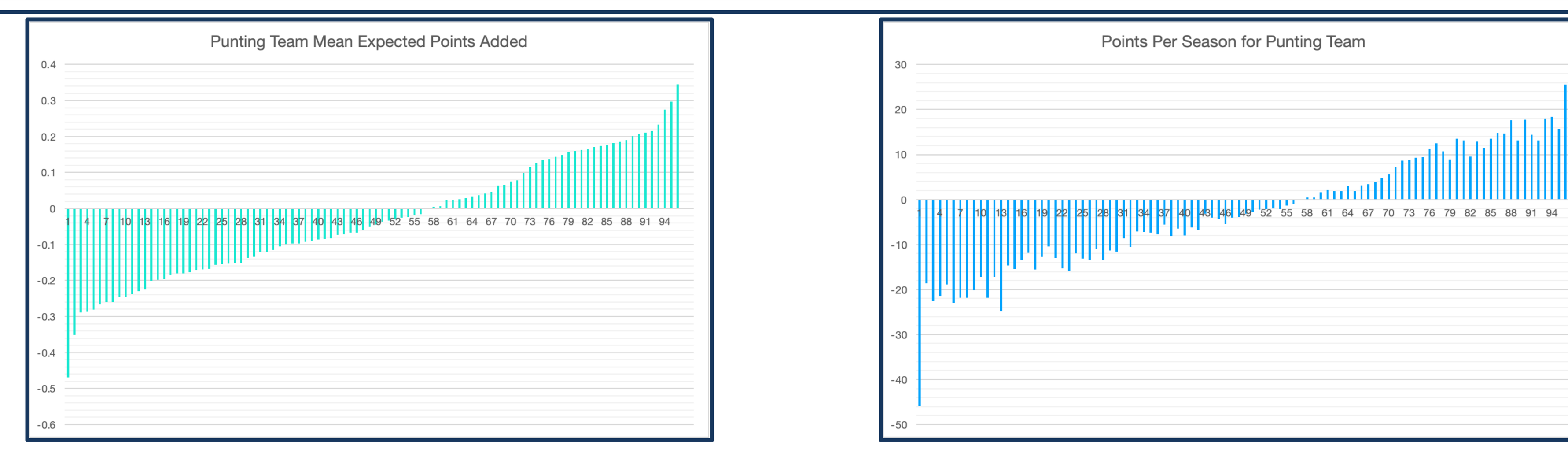


Figure 3—- Points Added to or Lost by Punting Team
The data set includes the average expected points added for every punting (offensive) NFL team from 2017-2019. The 32 teams were counted as 3 separate teams for each season, resulting in 96 teams. The total expected points per season resulting from a punt are also included for every team over the 3-year span. A negative value corresponds to the amount of expected points lost from the punting team's score by their punts while a positive value corresponds to the amount of expected points added to the punting team's score by their punts. For example, the Baltimore Ravens scored 468 points in 2020. The best punting teams (top 4 of 96) can benefit their team by upwards of 17 expected points per season (which is ~3.6% of the 2020 Ravens total points). The worst punting teams (8 of 96) cost their team upwards of 20 expected points per season (which is ~4.3% of the 2020 Ravens total points). Having an average punt team has a very small negative impact on the punting team over the course of a season, whereas a bad punt team has a significant impact on the punting team's total points over the course of a season.

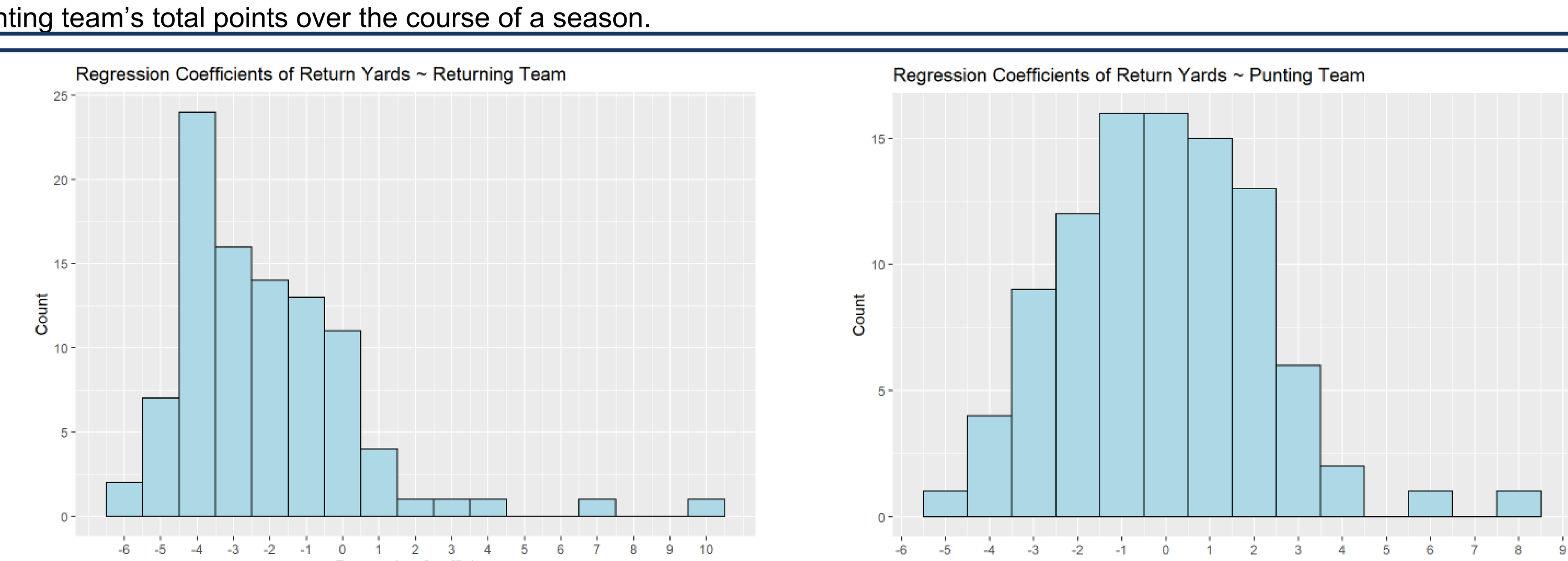
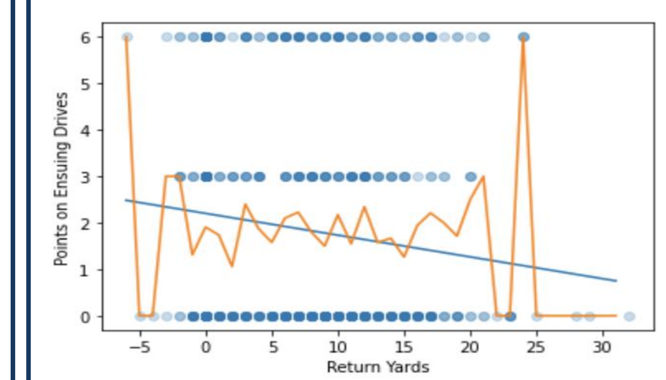


Figure 4—Using linear regression to predict length of punt return based on team alone
The data set spanned 3 years and 32 teams so there are 96 unique punting and return units. Punt return yards were regressed against both the punting units and return units. The Ravens' 2017 punt unit and return unit were used as baselines. Thus, the coefficients correspond to the expected difference in punt return length between the given unit and the 2017 Ravens unit. Despite the range in coefficients only a few were statistically significant. 3 return units were significantly worse and 1 was significantly better than the Ravens' 2017 return unit. 2 punt units were significantly worse, and none were significantly better than the Ravens' 2017 punt unit. This suggests that there is little variance in a team's ability to create or stop long punt returns.

−0.04679331084992585 2.201625897423842 −0.3601922514276086 0.026326805592008465 0.020198683045223032



0.0009233316466407724 1.9690884731158815 0.012871163318989916 0.9513087527431614 0.014956846098395878

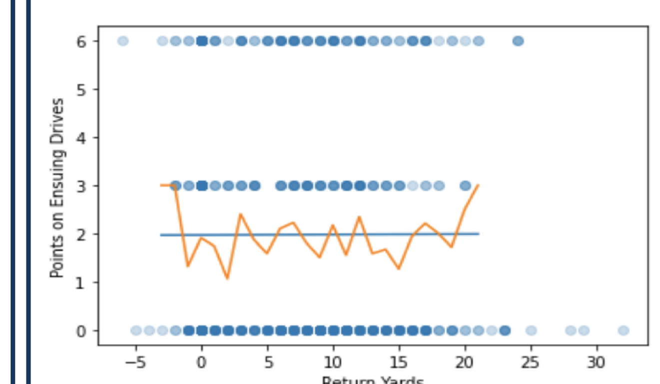


Figure 5 – Average Points on Drives Following Punts Beginning Between the 30 and 40 Yard Lines of the Return Team; First is Unfiltered Data, Second is Constrained
These two graphs are based around the idea of trying to prove/disprove momentum due to punt returns, specifically long ones. In this experiment, we took the average points of drives starting between the 30 and 40 yard line and compared those average points to the length of the return that preceded that drive. We ran this experiment on different yardage buckets as well. Theoretically, if momentum from punt returns existed, we should see an upward slope because teams score more points from the momentum. The second graph decreased the possible return yards observed to handle some sample size issues. The statistics above each graph are the following metrics from a linear regression: slope, intercept, r-value, p-value, and standard error.

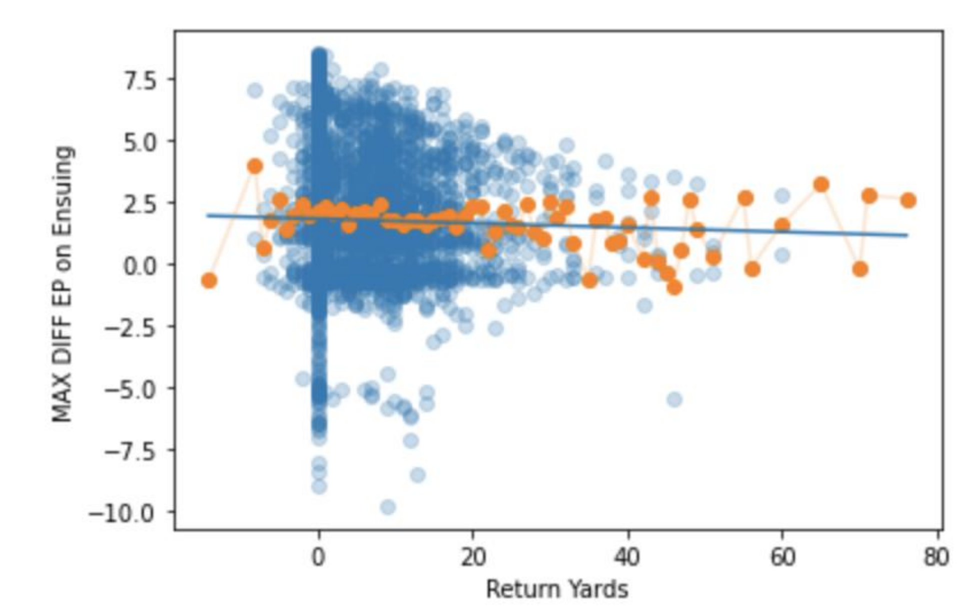


Figure 6 – Max Differential in Expected Points on Ensuing Drives vs. Punt Return Yards
We further attempted to observe momentum from punt returns by using an alternative method of drive scoring, instead of the more discrete metric of points scored: Maximum Differential, any time in the drive, in Expected Points from the initial on the drive. The advantage of taking the difference in expected points from the initial as a metric of the drive score is that it is a continuous metric. Furthermore, the potential advantage of using the maximum differential from the initial at any point in the drive is that it should capture the opportunity a drive creates, agnostic of potentially random events that would alter points scored on the drive (ex. Fumble at the opponent's 1-yard line). Regression on this metric would ideally show upward trend if momentum existed when compared to return yards. However, note that this metric is not perfect due to confounding factors like the potential for higher maximum differential when return yards are lower.

Conclusion

Even when accounting for in game variables the length of a punt return remains highly variable. Punting and punt return units have a small impact on their team's scoring capabilities. Most of the punt and return units are comparable and are not significantly better or worse than others. Long punt returns do not create quantifiable momentum on the ensuing drive.