

---

---

# Quantifying Leverage at the Point of Attack

Big Data Bowl 2020

## 1 - Introduction

*“Why do NFL teams run so often?”*

As data available from NFL games increases, fans of the game are increasingly using a data-driven approach to evaluate strategies that teams have implemented for many years. One of the biggest points of contention is the need to “establish the run”. As of writing this paper the top 5 teams in rushing yards per game in 2019 are:

- 1) Baltimore Ravens
- 2) San Francisco 49ers
- 3) Seattle Seahawks
- 4) Minnesota Vikings
- 5) Dallas Cowboys

At least four of these teams will be advancing to the playoffs in 2019 while the fifth is still in strong contention to join the others. This is often cited as evidence to invest in a strong rushing attack. Critics will often counter that teams who kneel the ball often also seem to have strong records, implying that there is a bit of confounding going on with both situations. The evidence against the necessity of a strong running game seems to have increased with the accessibility of expected points added (EPA<sup>1</sup>) calculations for each NFL play through the R package nflscrapR [1]. It was found that rushing plays have a negative EPA on average [2]! While the result is certainly interesting, making inferences about individual rushing plays or team rushing attacks can be fall under ecological fallacies. Baltimore currently has the highest rushing EPA per play; should they rush more? Baltimore also has a higher average EPA when they pass than when they run; should they pass more? Is the strength of the run game helping the pass game or is it vice-versa?

There is no shortage of good questions and not a lot of metrics specifically tailored to running plays. Quarterbacks can be evaluated through traditional metrics such as passer rating or completion percentage; they can also be evaluated through advanced stats such as air yards [3] and QBR [4]. Unfortunately, such a wide set of tools is not available to evaluate run plays. Most attempts to evaluate blockers in rushing plays must rely on allocating credit of a rushing outcome to the lineman. Given the limitations of the data, this may be the best that is possible. However, with the expansion of tracking data in the NFL it should be possible to improve on these metrics

---

<sup>1</sup> Expected points quantify how many points a team should score on average, given their present situation. EPA calculates the expected points gained, or lost, on a play.

---

---

---

and decouple running back and lineman contributions to run plays. The purpose of this paper is to contribute to that task.

We present two metrics which will help evaluate the quality of a block at a given point in time:

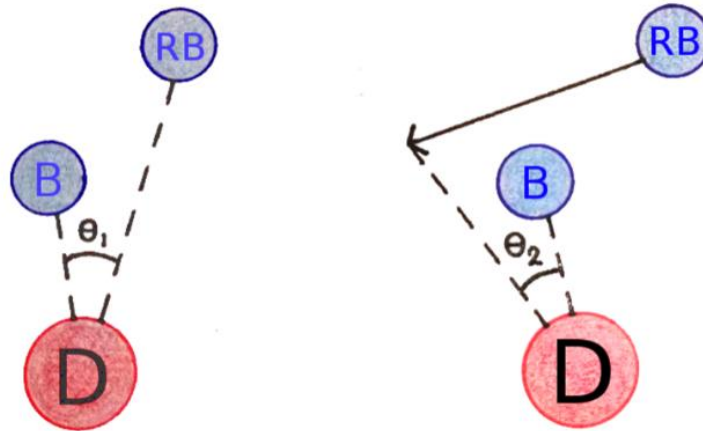
- Defender leverage – Measures the quality of the position a defender has, to make a tackle
- Blocker leverage – Measures the effectiveness of the blocker’s position

In section 2 we will define and describe these two metrics. In section 3 we will get to see an example of how these metrics evolve over a rush play and how they relate to that play’s outcome. Section 4 will give a holistic sense of these metrics and their approximate distributions. In section 5, we will use these metrics to evaluate and rank defenders through a 6-week span. Lastly, in sections 6 and 7, we will discuss some of the implication and future directions of both metrics. All results presented here are based on NFL tracking data from weeks 1-6 of the 2017 season, provided as part of the 2019 Big Data Bowl

## 2 - Leverages

Line coaches from youth football to the NFL stress the concept of leverage to their players. In this context we can define leverage as the physical advantage given by position. Lineman must focus on leverage of their body, leverage of their hands, as well as leverage at the point of attack. The first two would be difficult to quantify even with tracking data so we focus on the third of these. An old coaching adage describes leverage at the point of attack as “getting your butt between the defender and the ball-carrier”. If the butt is in fact between the defender and the ball-carrier, then the angle created by drawing lines from ball-carrier to defender and defender to blocker will be small. This type of angle underpins our metrics, but it is not sufficient to evaluate the quality of a block. To incorporate the dynamic nature of run plays and how angles between players can have different interpretations over time, we characterized two different types of leverage:

### Defender Leverage



Figures 1 and 2:  $\theta_1$  is the angle created by a blocker, the defender, and the running back (RB). A new angle,  $\theta_2$  is created if the RB's location is projected in time.

In Fig. 1 we see a ball-carrier, a defender, and a blocker. Does this defender have good leverage to make a play on the ball-carrier? Of course, it depends on where the point of attack is. If the ball-carrier is running straight down the middle, then he should run straight into the defender. However, if the ball-carrier runs to the left behind his blocker, then the defender will run into the blocker if he continues to pursue the ball-carrier; the blocker would then be in a good position to make the block and the defender would be in a poor position to make the tackle. For this reason, in order to evaluate a defender's position, we need to anticipate the ball-carrier's point of attack.

At each frame, the tracking data contains the direction each player is going as well as their speed. From this we can estimate each player's position in the future, however the estimate will become increasingly unreliable the farther out its projected. For this reason, we only projected the ball-carrier's future position 0.5 seconds from each frame. Using this projected point, we define two important angles:

- $\theta_1$ : The angle created by joining  $RB$  to  $D$  and  $D$  to  $B$
- $\theta_2$ : The angle created by joining  $RB^+$  to  $D$  and  $D$  to  $B$

Where:

- $RB$ : The ball-carrier's current location
- $RB^+$ : The ball-carrier's projected location in 0.5 seconds
- $B$ : The blocker's current location
- $D$ : The defender's current location

If  $\theta_2$  is smaller than  $\theta_1$  then that means that the blocker's position is becoming better as the running back moves. This is because, as stated earlier, a small angle between the players indicates that the blocker is in good position.

From these two angles we define defender leverage as:

$$\delta = \theta_2 - \theta_1$$

If the defender's position is improving as the running back approaches the point of attack, then  $\theta_2$  will be larger than  $\theta_1$  and  $\delta$  will be positive. Conversely, if his position is becoming worse, then  $\theta_1$  will be larger than  $\theta_2$  and  $\delta$  will be negative.

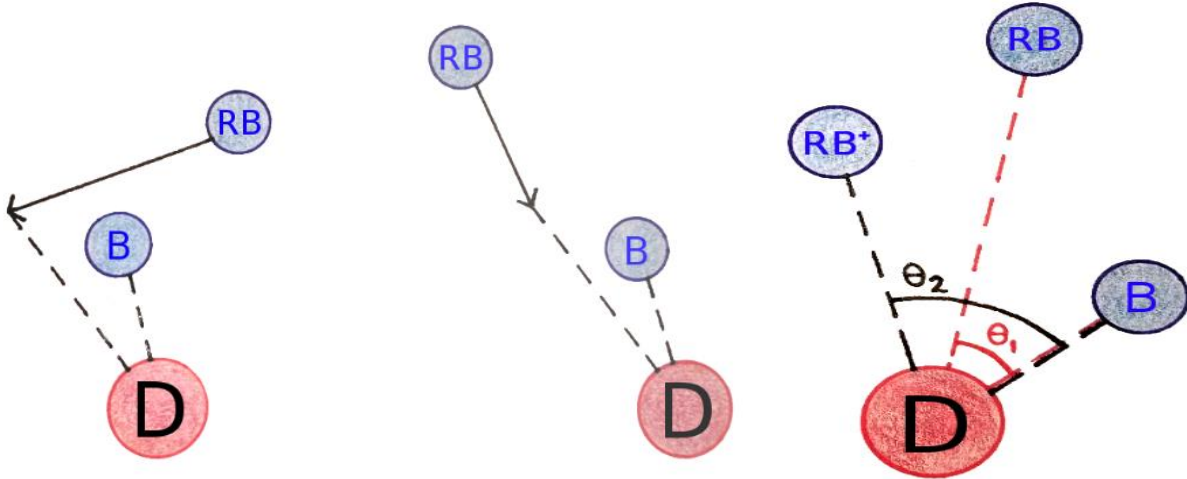


Figure 3 and 4: Both diagrams have the same  $\theta_2$  values, yet the blocking is clearly more advantageous in figure 3.

Figure 5:  $\delta = \theta_2 - \theta_1$  can alternatively viewed as the angle between  $RB^+ - D - RB$ , multiplied by  $\pm 1$ . The sign is determined by the blocker's position within this angle.

If the ball-carrier's point of attack is what is important in determining leverage, it is reasonable to ask why we consider  $\theta_1$  at all and just use  $\theta_2$  instead of  $\delta$ . The difference between  $\theta_2$  and  $\theta_1$  can be thought of as how the important angles between players is changing over time. This contains important information about relative positioning that  $\theta_2$  cannot capture alone. Figs. 3 and 4 show situations in which the important angle at the point of attack,  $\theta_2$ , is the same for both plays but the situation for the defender is quite different. In Fig 5, the running back is heading straight for the defender. While the blocker should be able to step between the ball-carrier and the defender, the ball-carrier will still have the difficult choice of choosing a side of the blocker to run down. Alternatively, the trajectory in Fig. 4 shows a clear hole for the running back to attack, with the blocker only needing a step to block that path. In the former case, the important angle has shrunk, with the defender having a good position;  $\delta$  will be positive here. In the latter case  $\theta_2$  is smaller than  $\theta_1$  which indicates a negative  $\delta$ . As we would hope to convey, the defender's position is better in Fig. 5 than in Fig.4.

### Blocker Leverage

It might be unclear why  $\delta$  is called "defender" leverage when it also involves the blocker. In cases where leverage calculations are relevant and the blocker is in between the defender and the point of attack,  $\theta_2 - \theta_1$  can also be described as the angle between the ball-carrier, the defender, and where the ball-carrier is projected ( $RB - D - RB^+$ ) with a subtle difference: The sign of the angle is negative if the blocker is in that angle and positive otherwise. In other words, that angle

is positive if the defender has good leverage to get to the ball carrier. Even though  $\theta_2$  and  $\theta_1$  both involve the blocker, by taking the difference, his position only effects the sign of  $\delta$ . Therefore  $\delta$  is used for a defender's leverage. This result is illustrated in Fig 5.

We still want to evaluate the quality of the blockers position and use it in tandem with the defender's leverage to evaluate a block.

We define the blocker's leverage as:

- $\beta$ : The angle created by joining  $B$  to  $RB^+$  and  $RB^+$  to  $D$

The logic for defining it this way has been laid out previously. If the blocker is in a good position with respect to where the ball carrier is heading, then this angle will be small.

### 3 - Leverages in Action

Both blocker and defender leverages are measured at an instantaneous point in time. With the tracking data we can calculate leverages at each time frame and see how leverages evolve over a play. To see how this in action we will revisit game in which the Green Bay Packers thoroughly beat down the Chicago Bears in 2017.

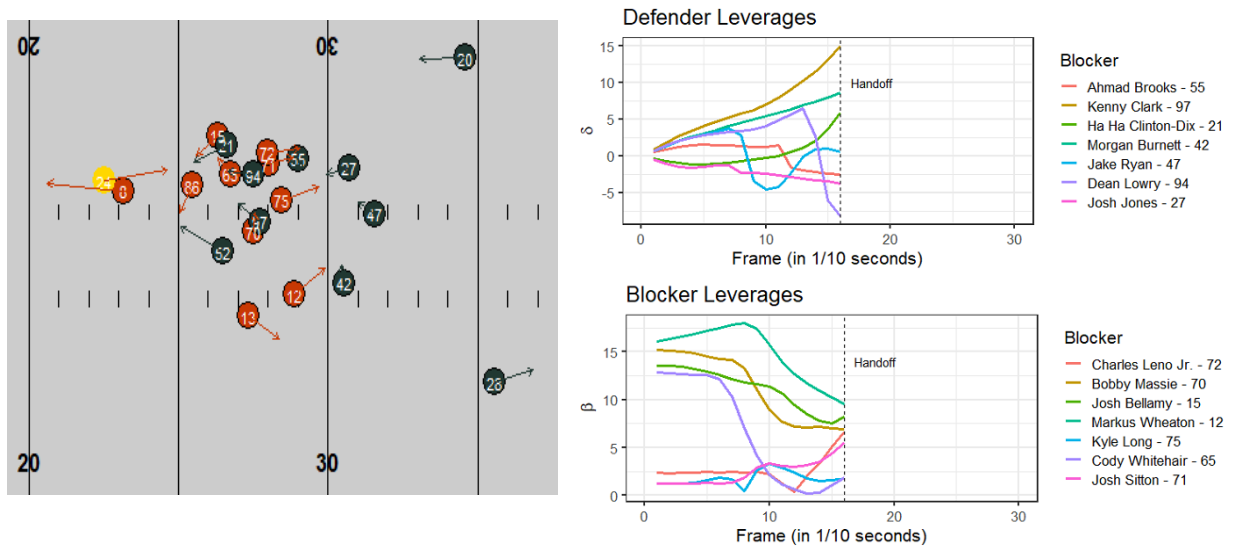


Figure 6: Each player's position and direction at the point of the handoff. The Chicago Bears are brown while the Green Bay Packers are a dark green. The ball-carrier is gold. It should be noted that there is an additional player down the field that has been cropped for viewability.

Figures 7 and 8: Defender leverages ( $\delta$ ) and blocker leverages ( $\beta$ ) for each engaged player. Players who are not engaged in a block over the course of the play will not have leverages calculated for them.

Fig 6. shows a first down running play at the point of the handoff. We can see that the defense has made a lot of penetration from the line of scrimmage located at the 28-yard line. Ideally for the offense, the running back (Jordan Howard - 24) will find a way to bounce to the outside as it

looks like his prospects along his current trajectory in the middle are not too favorable. The plot of leverages in Fig 7. can give similar information. The defenders with the highest leverage at the handoff have numbers 42, 97, and 21. Unfortunately for Howard, two of those defenders are close to his point of attack. Already it seems that this play is unlikely to succeed. On the other hand, 94 also is in close proximity to Howard but has one of the poorer leverages. At this point, none of the blockers have terrible blocking leverage  $\beta$  and most have been able to improve their positioning since the snap.

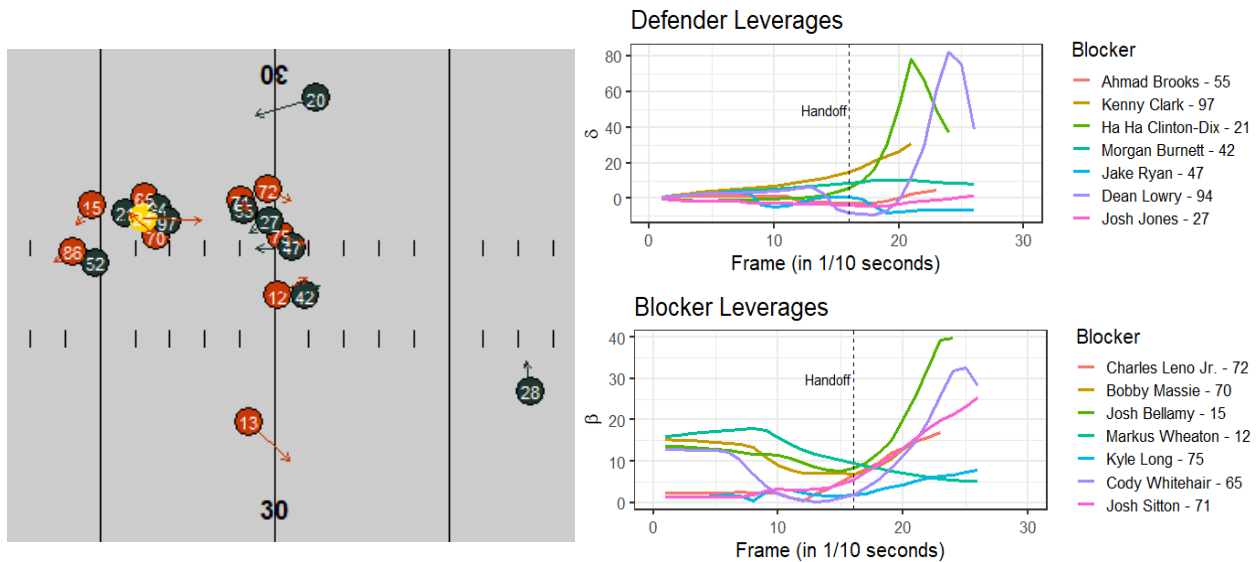


Figure 9: Each player's position and direction one second after the handoff.

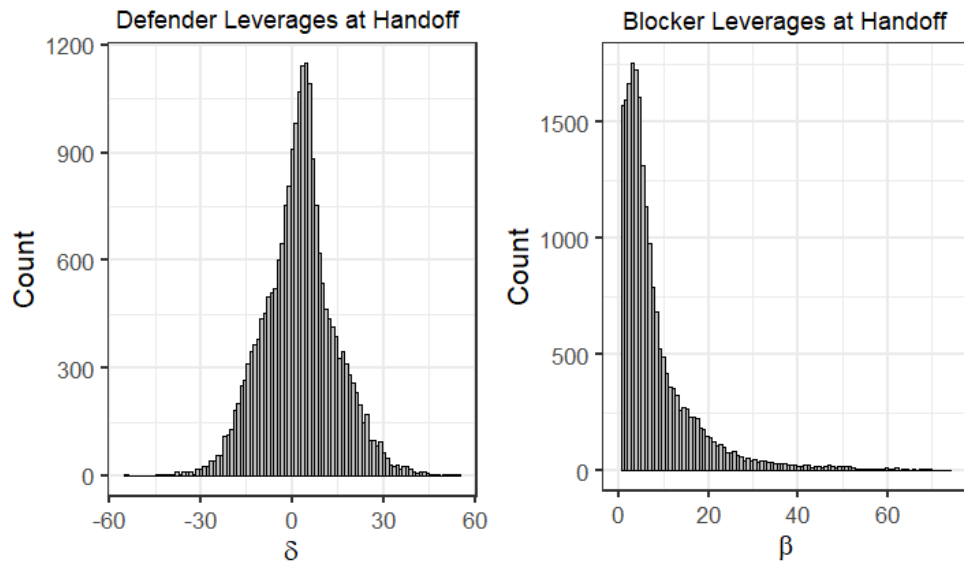
Figures 10 and 11: Defender leverages ( $\delta$ ) and blocker leverages ( $\beta$ ) for each engaged player displayed up to one second after the handoff.

If we examine one second after the handoff, we can see that Howard did not in fact bounce the play to the outside and he will be tackled imminently barring any Beast-Mode impressions. The three closest defenders who are about to make the tackle also have the highest  $\delta$ . It's interesting to see that 94 had one of the worst positions at the point of the handoff but was able to achieve excellent leverage in that second to help make the tackle. Conversely, 42 was never able to build on his leverage and 12 was able to make up the ground to block him. A lot can change in a second. Over that second, the blocking leverages  $\beta$  began to deteriorate for many blockers as well.

It is worth noting here that the time plots of leverages will have different sizes. Both leverages measure the quality of a block while the blocker and the defender are engaged. Once the defender beats their blocker, then the angles that inform  $\delta$  no longer have the same interpretation. Because of this, we stop calculating leverage when the defender beats their blocker.

---

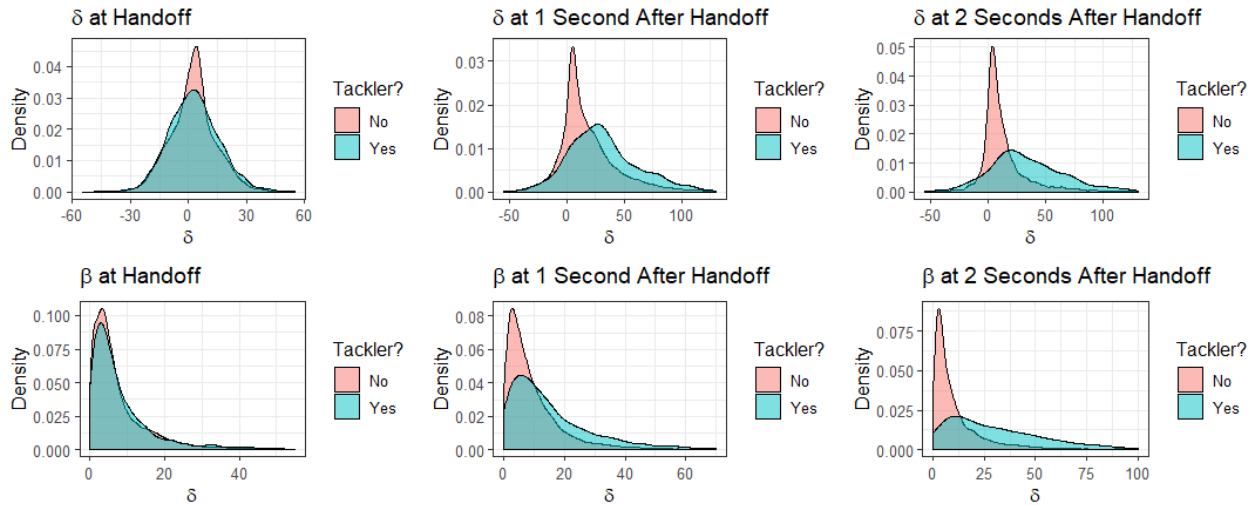
## 4 - Leverage Distributions



Figures 11 and 12: Histograms of leverage values collected at the handoff.

Based on the data we had, we were able to collect leverages across 4164 plays that occurred in 2017 from weeks 1 to 6. Figs. 11 and 12 show histograms of leverages collected at the handoff. The distribution of  $\delta$  has a slight skew to the positive side but still has a roughly symmetric shape. This indicates that, without consideration of their position, the defender has a slightly higher chance of being in a good position to make a play on the ball carrier.  $\beta$ 's distribution resembles an exponential distribution as it is skewed right with a light tail that diminishes around 50. From these plots, we can see what good and bad leverage looks like in a relative sense.

Even though  $\delta$  and  $\beta$  have physical interpretations, it can be difficult to put them into context without relating them to previously interpretable metrics. The obvious metric is the defense's goal each time the offense rushes the ball: a tackle. With this in mind, we can investigate how the leverage distributions change for those who make the tackle and those who do not. For  $\beta$ , we identify whether or not the defender they were blocking made the tackle. To simplify calculations, we only recorded one tackler per play although with a little bit more time we can account for group tackles. This improvement should further clarify the relationship between leverage and tackles as currently we are underestimating tackles. In some sense the tackles we are recording can be viewed as solo tackles.



Figures 13-15: Defensive leverage ( $\delta$ ) distributions of players who made the tackle on the play and those that did not. These distributions are shown at the handoff, one second after, and two seconds after.

Figures 16-18: Blocker leverage ( $\beta$ ) distributions of players based on whether the defender they were blocking made the tackle on the play and those that did not. These distributions are shown at the handoff, one second after, and two seconds after.

From Figs. 13 and 16, the leverage distribution of those who will make the tackle and those that will not is almost indistinguishable for both  $\delta$  and  $\beta$  at the point of the handoff. However, both begin to diverge as the leverages are evaluated farther away from the handoff. At two seconds, the distributions of  $\delta$  and  $\beta$  for the tacklers begin to have very heavy tails. For  $\delta$ , this means that the tackler is getting increasingly good position over the blocker to make a tackle; for  $\beta$ , this means that the blocker is losing their position against the defender. It is interesting to see that even for the non-tacklers, the distribution of  $\delta$  is still skewing right over time. It appears that it can be difficult for blockers to “get their butts” in front of where the ball carrier is going as a play progresses. Still, it seems that good leverage is not wholly sufficient to make a tackle.

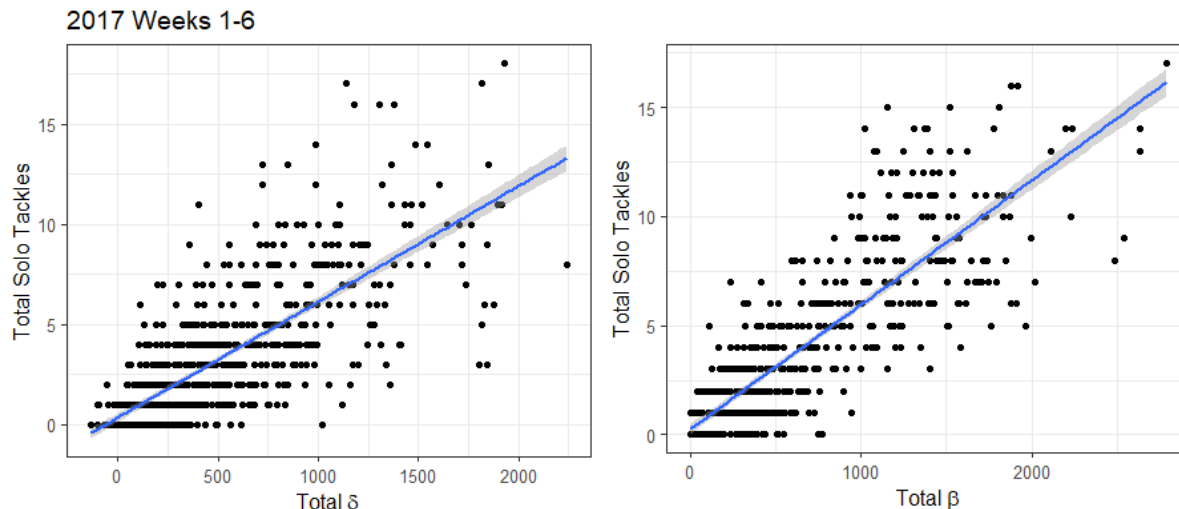
## 5 - Evaluating Defenders using Leverage

Even though good blocker leverage does not appear to be sufficient to make a tackle, it seems necessary. Another way to evaluate leverages in this time series context is by looking at the end of the time series. As alluded to before, we only measure leverages when they are relevant to a blocker engaged with a defender. For this reason, we stop evaluating leverage when either the defender beats the blocker or when the ball-carrier passes the engaged pair. The distributions of  $\delta$  and  $\beta$  at the end of each player’s respective time series look very similar to the plots in Figs. 15 and 18 after two seconds have passed from the handoff. This makes sense as this is likely the time a play will end or at the very least the ball-carrier will pass the engaged pair. When the time series ends, the  $\delta$  of tacklers has only a bit of the distribution tail in the negative values.

Therefore, defenders who constantly achieve high  $\delta$  values are more likely to make tackles. Over the 6-week period in the study, each defender’s  $\delta$  and each blocker’s  $\beta$  were recorded at the end



of each of their respective leverage time series as well as whether the defender (or the person the blocker was blocking) made the tackle. We can see from Fig that there is a strong relationship between both total  $\beta$  and total  $\delta$  with the amount of tackles achieved by that player over the 6-week span.



Figures 19-20: Total defensive leverage accumulated and total solo tackles for each defensive player in weeks 1-6. A similar plot is shown for blocking leverage, but the tackles refer to tackles achieved by those that the blocker was blocking.

The Spearman correlation [5] between total  $\delta$  and total tackles is 0.79 and the same correlation between total  $\beta$  and total tackles is 0.87. Clearly both of these metrics have a strong relationship with tackles.

One of the goals of most advanced metrics is to account for chance results that affect an observed outcome but may not be predictive in the future. For example, if a quarterback throws a ball right into a defender’s hands but it bounces off of them into the receiver’s hands for a touchdown, then it shows up as a great play on the box score. However Pro Football Focus’ (PFF) metrics [6] will not grade this play highly for the quarterback since the play should have resulted in an interception. Tackles as a metric can fall victim to similar effects in a play. A defender can have great position but be held by a blocker. If the ref doesn’t see it then the defender can’t be credited with having good position. Also, the defender can get off the block but fall victim to a great move by the running back.  $\delta$  and  $\beta$  can potentially account for these things.

Over the course of the first 6 weeks of the 2017 season, the players who racked up the most defensive leverage at the end of each of their respective leverage time series is given in Table 1. Of these 10, 3 were also in the top 10 in total solo tackles during that span. Of the players on this list that were not in the top 10 in solo tackles, only the bottom 3 fell out of the top 50. Unfortunately, it is difficult to know how predictive  $\delta$  is of future tackles without collecting more data. Based on the amount of leverage he was consistently creating, it’s hard to believe that Karlos Dansby would not be due for an uptick in solo tackles in the future...I don’t know many people who will claim that Dansby is a poor tackler.

---

Table 1: Defensive leverage rankings through weeks 1-6 2017

Player	Team	Total $\delta$	Solo Tackles
Brandon Mebane	LAC	2239	8
Linval Joseph	MIN	1928	18
Mark Barron	LAR	1914	11
Jason Pierre-Paul	NYG	1899	11
Arby Jones	JAC	1877	6
Blake Martinez	GB	1852	13
Darron Lee	NYJ	1843	9
Karlos Dansby	ARI	1842	3
Hayes Pullard	CLE	1832	6
Akeim Hicks	CHI	1818	17

## 6 - Discussion

A similar evaluation of blockers could be performed using  $\beta$ . This is certainly an appealing thing to do as it is even more correlated with the solo tackles recorded than  $\delta$ ; we would caution against this at the moment. The blocker's position is a function of the blocker but it is also a function of the running back and which way he is going. A blocker can be in the correct position only for a running back to choose the incorrect hole and suddenly the blocker's  $\beta$  could be large. It doesn't seem fair to lineman to evaluate their skill on a metric that we can't decouple the running back's ability from the lineman's. This metric is still clearly useful and is a ground for future study. If the ball carrier's decision making can be viewed through the lens of  $\beta$  (among other things) and decoupled from the lineman's current position, then it could be possible to make some headway in determining the ball-carriers value to a rushing attack.

On the other hand,  $\delta$  seems like a reasonable way to evaluate defenders on running plays. Even though ball-carriers and blockers are not decoupled, it really doesn't matter much to the defender. Can he make it into a spot where the running back wants to go? That's all that matters from a leverage perspective. Immediate future work in this area will consist of better classification of group tackles. Since we are not taking this into account there are likely quite a bit of defenders in the tails of the  $\delta$  distribution of the non-tacklers who belong in the tackler distribution. Maybe that is where all of Dansby's tackles are.

It has been shown that both  $\delta$  and  $\beta$  are useful metrics for describing a run play as it progresses. The further along a play is, it is very likely that the tackler has high leverage. With this in mind, it is interesting to see that there is very little difference in the leverage distributions between those who will make the tackle and those who will not at the handoff. This implies that there is a lot of information missing at the handoff that will be relevant to the play outcome. For this reason, it is risky to make strong generalizations based on the Big Data Bowl 2020 data that was released. This dataset contained far more plays (all regular season games from 2017-2018) but only contained the tracking data at the handoff. While impressive predictive feats can be

---

---

---

accomplished with this data, inferential claims based on this data could be misplaced if so much can change within even just a second of the handoff.

## 7 - Conclusion

Evaluating run plays is difficult. Outside of traditional metrics such as rushing yards and tackles there doesn't seem to be many tools for data-driven analysts. As the use of the NFL's tracking data expands, it will hold the key to providing insights on which themes create an effective rushing offense (or which ones are useful in stopping a rushing attack). The concept of point of attack leverage is something that any lineman is aware of and is of a lot of use to him. We have presented two novel metrics for evaluating the quality of this leverage for any pair of engaged players at any point in time. It has been shown that the metrics represent real-time observable positional traits that occur during a run play; these traits are highly correlated with intuitive metrics such as tackles yet still provide additional meaningful information.

What we believe is more exciting about  $\delta$  and  $\beta$  is that there is plenty of ways they can be improved upon through future work. With the NFL's full tracking dataset, both metrics can be tested to see how predictive they are of future tackles; this is something that would be valuable to talent evaluators looking for an edge in acquiring run-stoppers. Also  $\beta$ , and how it changes over time, could provide some valuable insights into running back decision making. Hopefully, leverages  $\delta$  and  $\beta$  can help lay some of the groundwork in a data-driven understanding of running plays in the NFL.

---

---

## 8 – Acknowledgements

Thank you to Saga Tuitele for crucial insights on how blocking experts think about running plays and Cat Wright for another year of award-winning illustrations.

## 9 – References

- [1] Maksim Horowitz, Ron Yurko, and Samuel Ventura. nflscrapR: Compiling the NFL Play-by-Play API for easy use in R, 2018. URL <https://github.com/maksimhorowitz/nflscrapR>. R package version 1.8.1.
- [2] Melin, J. (2018, November 3). The Pass and the Pass Nots: Sports Analytics Group at Berkeley. Retrieved from <https://sportsanalytics.berkeley.edu/articles/pass-and-passnotes.html>.
- [3] NGS: NFL Next Gen Stats. (n.d.). Retrieved from <https://nextgenstats.nfl.com/glossary>.
- [4] Oliver, D. (2011, August 4). Guide to the Total Quarterback Rating. Retrieved from [https://www.espn.com/nfl/story/\\_/id/6833215/explaining-statistics-total-quarterback-rating](https://www.espn.com/nfl/story/_/id/6833215/explaining-statistics-total-quarterback-rating).
- [5] Myers, L., & Sirois, M. J. (2004). Spearman correlation coefficients, differences between. Encyclopedia of statistical sciences, 12.
- [6] PFF Player Grades. (n.d.). Retrieved from <https://www.pff.com/pff-player-grades>.