



PFF WAR: Modeling Player Value in American Football

Submission ID: 1548728

Track: Other Sports

Player valuation is one of the most important problems in all of team sports. In this paper we use Pro Football Focus (PFF) player grades and participation data to develop a wins above replacement (WAR) model for the National Football League. We find that PFF WAR at the player level is more stable than traditional, or even advanced, measures of performance, and yields dramatic insights into the value of different positions. The number of wins a team accrues through its players' WAR is more stable than other means of measuring team success, such as Pythagorean win totals. We finish the paper with a discussion of the value of each position in the NFL draft and add nuance to the research suggesting that trading up in the draft is a negative-expected-value move.

1. Introduction

Player valuation is one of the most important problems in all of team sports. While this problem has been addressed at various levels of satisfaction in baseball [1], [23], basketball [2] and hockey [24], [6], [13], it has largely been unsolved in football, due to unavailable data for positions like offensive line and substantial variations in the relative values of different positions. Yurko et al. [25] used publicly available play-by-play data to estimate wins above replacement (WAR) values for quarterbacks, receivers and running backs. Estimates for punters [4] have also been computed, and the most-recent work of ESPN Sports Analytics [9] has modified the +/- approach in basketball to college football beginning in the fall of 2019. Pro Football Reference's [20] Approximate Value (AV) is the most widely used open-source measure for player value, and although the currency is not "wins", it is a very useful measure by which to compare players across positions and seasons.

Pro Football Focus (PFF) [19] offers hope in the area of player valuation due to its unique player grades, which provide play-by-play participation and performance data for each player on the football field, from the quarterback to the gunner on the punt team. These grades have been transformational in the football community - from appearing on NBC's Sunday Night Football to playing a part in contract negotiations for players and teams. The PFF grades have become the go-to place for establishing *how good* a player is at playing football. An open question remained as to how well it could establish *how valuable* a player is. A quarterback with a 67 PFF grade as a passer is not worth the same to his team as a running back with a 67 grade as a rusher. Furthermore, a left tackle with a 67 grade in 100 snaps is not worth the same to his team as a left tackle with the exact same grade in 1000 snaps.

Therefore, a player valuation model needs to incorporate a) how well the player played b) what the player did and how important is doing that well, and c) how often did the player do the various things he did. In this paper we introduce PFF Wins Above Replacement (PFF WAR). PFF WAR is computed using PFF's player grades and participation data in conjunction with a Massey matrix model [15] infused with machine learning [12] to determine how important each facet of play is in football. We find that, on balance, this WAR metric is very stable season-to-season during the PFF



era (2006-present), although this stability varies by position group. Defensive backs and wide receivers are the most valuable non-quarterbacks on average, but the former's WAR estimates are far less stable than those from other positions on the defense.

We find that the total team wins implied by WAR are more stable than actual wins earned and Pythagorean-implied win totals [11] and that the total number of wins above replacement generated during a rookie contract falls off monotonically with draft position. However, differences between quarterbacks and non-quarterbacks are significant enough for both teams to win a trade involving draft picks.

2. The Massey Rating System

2.1. Original Formulation

The Massey rating system [15] is a matrix-based system of evaluating the overall performance of teams in a league by assuming that the overall strength of a team is a product of a) the interactions between teams in a given league and b) the existing strengths of the teams with which a given team interacts. Modeling a) is done by the construction of the *Massey matrix*, which is an n by n matrix, where n is the number of teams in the given league. The i th term along the main diagonal of the Massey matrix is the total number of games played by the n th team in the league during a given period of interest (usually one season). The i,j th element of the matrix is the number of games played between team i and team j during that same period, multiplied by -1 . It's clear that every Massey matrix M is a symmetry matrix, and each of the columns of the Massey matrix sum to zero.

Modeling b) is done by using an n -dimensional vector of team strengths over the period of interest. In the original Massey framework, this amounted to the total net points for that team, which was positive if that team scored more points than its opponents during the season and negative if they scored fewer. It's clear that the vector of team strengths f sums to zero.

The resulting rating vector, r , for which the i th element is the overall rating for the i th team in the league during the period of interest, solves the matrix-vector equation:

$$Mr = f, \tag{1}$$

which represents the assumption that each team's strength is a linear combination of the difference between its strength and those of each of its opponents. Offensive and defensive ratings can be extracted from r in a straightforward way by using the positive and negative parts of f .

2.2. Adjusting Massey Ratings Using PFF Grades

The Massey rating system is lacking in some contextual places. For example, a team with a substantial number of defensive touchdowns or touchdowns generated by the offense but from the aid of a short field will actually have a higher offensive Massey rating than they should, while a team whose offense constantly puts them in short fields or gives up points themselves will have a lower Massey defensive rating than they should. Special teams touchdowns scored are counted on the offensive side, when often they are very much influenced by a team's defense.



Additionally, football is a noisy game to begin with, where passes that should be intercepted can bounce off a defender's hands into those of a receiver for an offensive touchdown. Perfect passes are dropped, and fumble luck persists for an entire season in many cases. These issues can, in many ways, be addressed with PFF grades. In the PFF system each player is assigned a grade between -2 and 2, in increments of 0.5, and this grade can be classified as passing, rushing, receiving, run blocking, pass blocking, screen blocking, run defense, pass rushing or coverage, as well as a plethora of special teams designations. Each player earns a grade for avoiding penalties on offense, defense and special teams as well. An example of a +0.5 grade earned by a quarterback is an accurately-thrown completion traveling 10 yards on third down and nine yards for a first down, while an example of a -0.5 coverage grade is applied to the cornerback who was beat on said play for the reception (or even in the case the pass was dropped).

Thus, we can get around some the noise issues inherent in the Massey model through using PFF grades. While most positively-graded plays by quarterbacks will result in completions and/or touchdowns, and negatively-graded throws incompletions and/or interceptions, capture much of what score differential captures, the noisier plays will be given null (zero) grades, and plays that "should" have resulted in a different outcome will be graded accordingly, making the vector f that uses them part of a model that is more predictive than the original version.

To create f for our PFF Massey model, we first normalized the grades in each facet of play by season (a fixed effect) and season-long position (a random effect [17]) at the play level. We needed to normalize by the former because our grading system has evolved since its inception to comply with feedback from team clients and consultants. We controlled for latter because summary statistics within the same facet (e.g. pass blocking) were different between different positions (e.g. guard and tackle). After normalization, the average player at a position will earn a zero cumulative grade in each facet of play over the course of a season.

We then used a random forest model [12] to scale PFF facet grades by their variable importance relative to adjusted team wins in each season (including playoffs). Adjusted records are computed by giving a team credit for a full win or loss if the score differential was nine or more, and half a win and half of a loss otherwise. Some facets of play needed to be split up due to grades generated at different positions yielding different value relative to team wins (e.g. . The three most-important facets are passing, receiving by wide receivers and coverage by defensive backs, followed by the remaining non-special teams' facets (offensive penalty grade is more important than defensive penalty grade, which is consistent with the theme that offense is more important than defense [3]). Once the vector f is built, PFF Massey ratings are computed by solving an equation analogous to (1).

Metric	Cor(Metric year n+1)	Cor(Win Pct year n)	Cor(Win Pct year n+1)
PFF Massey total	0.45	0.73	0.35
PFF Massey offensive	0.46	0.67	0.33
PFF Massey defensive	0.37	0.50	0.25

Table 1: Pearson correlation [22] coefficients between PFF Massey ratings and PFF Massey ratings in subsequent years, win percentage in concurrent and subsequent years (including playoffs).

Some model diagnostics on the PFF Massey ratings are in Table 1. The second column in Table 1 shows what has been known for some time [3], that offense is what wins games in football more



than defense does. The third column shows that all three PFF Massey ratings offer some predictive power with respect to (unadjusted) total team wins in subsequent seasons, despite not accounting for changes in front office, coaches or players.

3. Player Valuation

3.1. Wins Above Average

The first step in a wins above replacement model is the derivation of a wins above average (WAA) model. The reason for this is that a replacement player is a difficult concept to define, but a team of replacement-level players is more straightforward. Furthermore, on the player level, the following equation holds:

$$\begin{aligned}
 WAR &= \text{player wins} - \text{replacement player wins} \\
 &= \text{player wins} - \text{average player wins} \\
 &\quad + \text{average player wins} - \text{replacement player wins}.
 \end{aligned}
 \tag{2}$$

The first two terms of the last expression are exactly WAA, which varies by player. Given a known fixed snap count for each facet, the final two terms above are fixed league-wide, making WAA the necessary entity to compute in this situation.

To compute WAA for a player, we compute the PFF Massey ratings for his team(s) twice - once with his grades present in the data set and once with them replaced by a 0 (average) - graded player with the same number of snaps in each facet of play. The wins above average for each player is then the difference in the number of wins implied (via a linear regression model) between the two PFF Massey ratings.

WAA estimates were computed for each player from 2006 to 2018, which amounts to over 25,500 player seasons. Year-to-year correlations in WAA for all positions (with no snap restriction) is 0.41, while players who played 250 or more snaps (on offense and/or defense) in season n (with no restriction on year $n + 1$) had a WAA that correlated at a rate of 0.48 (0.58 for quarterbacks).

3.2. Wins Above Replacement

To compute the last two components of (2), we need to define thresholds for a team of average players and a team full of replacement-level players. We assume a team of average players will win an average of eight games, while a team of replacement-level players will win an average of three games, leaving 171 total wins above replacement to allocate per season (including playoffs).

We assume that all 171 of these wins are found on either the offensive or defensive side of the ball (e.g. an average special teams player is a replacement-level player), and are apportioned according to the weights in the PFF Massey ratings on a per-snap basis for each facet. Each player's WAR is therefore his WAA value plus the sum of all his per-snap apportionment of the 171 total WAR multiplied by his number of snaps in each facet of play. Since $8 > 3$, a player's WAR will always be higher than his WAA, but the difference will be bigger for players that a) play more snaps and b) are utilized in more-important roles.



Table 2 contains the mean and standard deviation in WAR values for each position for players than play above the 60th percentile or more snaps in a given season at their position (for quarterbacks, this is just over 500 snaps), along with the year-to-year stability of WAR for each position. Notice that, as expected, quarterback is by far the most-important position, and as a measurement of quarterback play, WAR is the most-stable we've seen. For example, ESPN's QBR [10] is correlated year-to-year at a rate of 0.43, which is roughly the same as expected points added on passing plays (0.45, [25]). Both metrics are drastic improvements over traditional quarterback rating (0.37). None of these values rival that of PFF WAR during the same time period, though.

Wide receiver is the second most-valuable position, in many ways because they must capitalize on the process generated by quarterbacks in the form of accurate throws, but also can generate substantial value on expected, or even inaccurate, throws by making contested catches or generating yards after the catch. They can also lose their team significant value if they fail to catch accurate passes, and hence higher coefficient of variation in their WAR values relative to the quarterback position.

While the value of edge and interior players are similar on average, variation is higher at the edge position, since pressure (lack of pressure) from the edge is more valuable (detrimental) than pressure from the interior (Figure 1 (a) [7]). The more a player is in coverage, the more valuable they are, which is consistent with a conclusion we have that coverage is more valuable than pressure, all else being equal [8]. What earns the trench players a place back in the discussion is their substantial reliability in the way of year-to-year correlation in WAR, a virtue that players playing further away from the ball on defense cannot claim.

Position	n	Mean in WAR	Coefficient of Variation in WAR	Year-to-Year Correlation in WAR
QB	994	1.63	0.70	0.62
RB/FB	2373	0.10	0.64	0.53
WR	2864	0.28	0.84	0.52
TE	1621	0.18	0.62	0.66
T	1543	0.09	1.09	0.49
G	1604	0.10	1.11	0.57
C	708	0.10	1.08	0.50
DI	2559	0.06	1.34	0.68
ED	2259	0.06	1.54	0.61
LB	2721	0.11	0.83	0.51
CB	2733	0.23	0.91	0.29
S	2169	0.23	0.77	0.30

Table 2: Average WAR values, coefficient of variation in WAR values and year-to-year correlation in WAR values for each of the broad position categories in the National Football League. Abbreviations are as follows: Quarterback (QB), running back/fullback (RB/FB), wide receiver (WR), tight end (TE), tackle (T), guard (G), center (C), interior defensive player (DI), edge defender (ED), linebacker (LB), cornerback (CB), safety (S)

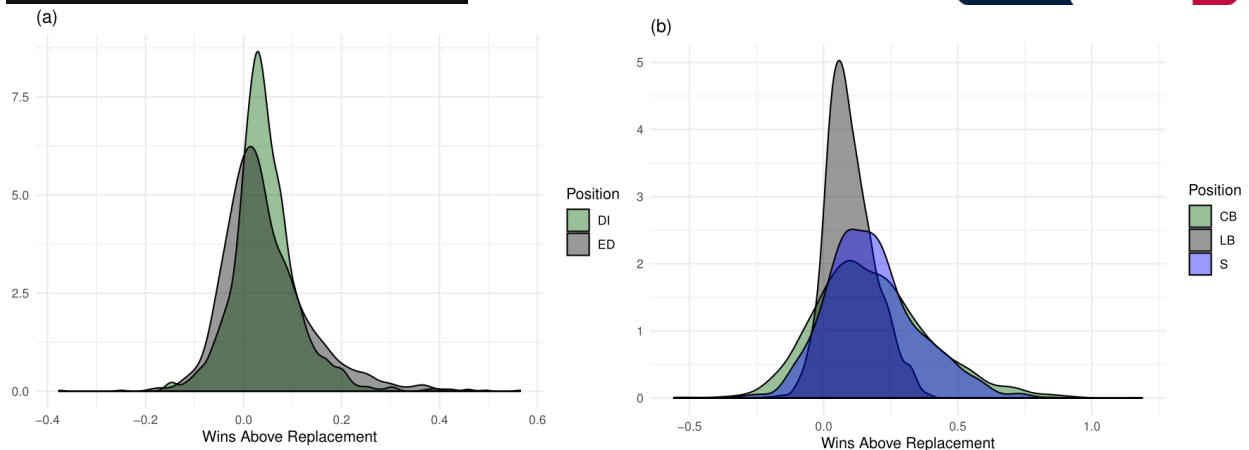


Figure 1: Distributions of seasonal WAR for positions along the defensive line (a) and the defensive back seven (b). Thresholds are 250 snaps for each position.

It is surprising that the average WAR values for each of the offensive line position designation are similar, but the similarity in average value between guard and tackle is due to the existence of a few guards generating outlier-level values, as tackles generate the most high-but-not-outlier values, while centers the most moderately-high ones (Figure 2 (a)). Running backs can add some value (even with their low snap counts) due in large part to their receiving. That said, it's telling that the sum of an average offensive line's WAR is roughly five-times that of an average running back, and an average wide receiver's is roughly three-times higher. Tight ends are both more valuable than running backs on average and more stable than any position group other than interior defensive linemen. As previously stated, wide receivers have the potential to add a great deal of value, but also to cost their team value due to their ability to turn positively-graded passes from their quarterback into negative plays for an offense due to drops or poorly run routes (Figure 2 (b)).

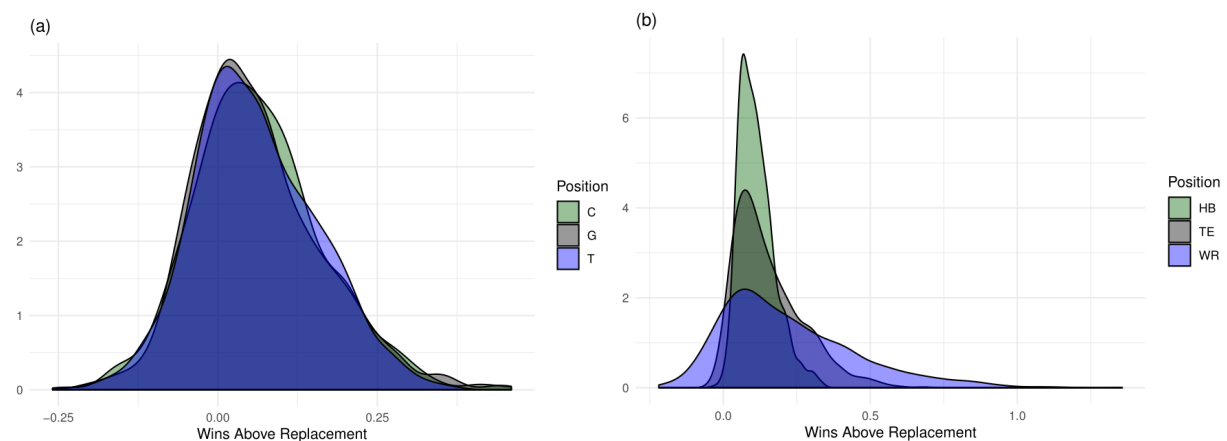


Figure 2: Distributions of seasonal WAR for positions along the offensive line (a) and the offensive skill positions (b). Thresholds are 250 snaps for each position.

Using every player in the sample and 250 as the snap threshold in the *first season of the year-to-year comparison*, WAR is correlated at a rate of 0.74 league-wide - which is better than almost every individual player metric or PFF player grade that we've studied at the NFL level. For example, Pro



Football Reference's [20] Approximate Value metric, the current standard singular number by which we evaluate performance of NFL players, is correlated at a rate of 0.64 since 1960, and 0.65 during the PFF era (2006-present).

4. Applications

4.1. Implied Team Wins

With WAR estimated for every player on every team since 2006, we examine how many wins a team "should" have had based on the quality of their players during a season by summing the entire team's WAR values and adding three wins. Much like analysts use Pythagorean implied win totals [11] to find over- and underachieving teams, analyzing the discrepancy between a team's actual wins and those that would be implied by aggregating its player's WAR can help us understand how lucky/unlucky the team was, how good/poor its coaching performed during the season, or where to look for possible places in which the whole is more or less than the sum of the parts in the game of football, so to speak.

Figure 3 (a) shows the distribution in the differences between actual wins (including playoffs) versus those implied by PFF WAR, which includes substantial variability. Some notable discrepancies include the 2006 Chicago Bears (8.3 implied wins versus 15 total wins) and the 2007 Green Bay Packers (8.8 versus 14). Both teams included quarterbacks that did not grade very well in our system in Rex Grossman and Brett Favre, and both failed to win even eight games the following season. The 2006 Detroit Lions (8.2 versus 3), 2017 Cleveland Browns (6.0 versus 0) were both teams whose players performed far better than their win-loss record (a finding consistent with Pro Football Reference's Expected W-L based on points for/against [20]). Both teams won seven games the following year. Teams in the seven-to-nine-win range that also win a playoff game will have a higher discrepancy by virtue of having played an extra two or more games, as will teams in the 10-11-win range that don't make the playoffs.

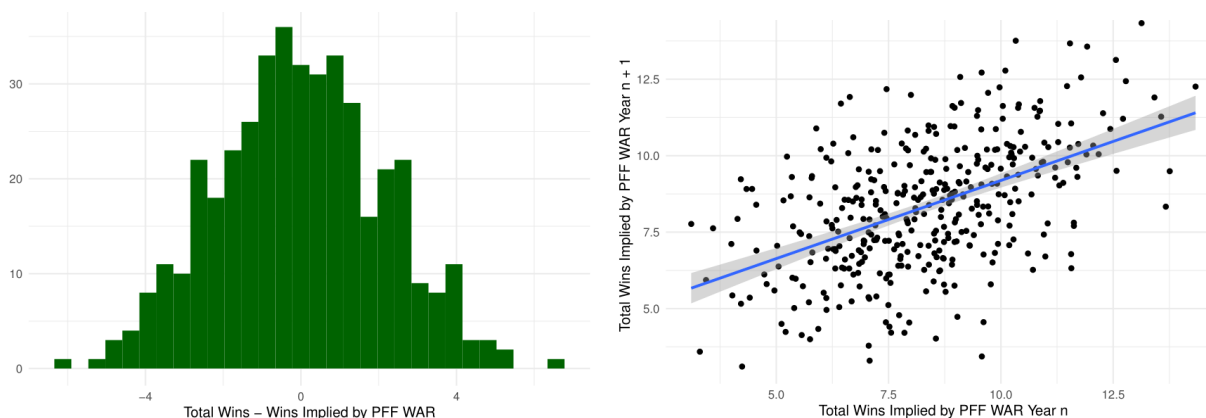


Figure 3: The differences between actual team wins (including playoffs) and wins implied by PFF WAR (a). Year-to-year stability in wins implied by PFF WAR, with correlation coefficient $r = 0.50$.

While it's difficult to draw any causal conclusions with respect to coaching efficacy, the Kansas City Chiefs have averaged more than two wins over those implied by WAR during head coach Andy Reid era (2013 - present), which is the most over that time period. Arizona was a top-five team during Bruce Arians' tenure with the Cardinals (0.97 wins above implied). It's a surprise to absolutely no



one that the Cleveland Browns have managed just under three wins per season under that implied by WAR during the entire PFF era (2006 - present).

Unsurprisingly, wins implied by PFF WAR are more stable season-to-season ($r = 0.50$, Figure 3 (b)) than actual wins (0.36, similar when one doesn't include postseason) or total team AV during that same time ($r = 0.42$). Wins implied by PFF WAR predict total wins the following season better than total wins do even without roster adjustments and adding roster adjustments (without incorporating projections for rookies) yields better predictive power than Pythagorean win totals do ($r = 0.43$ versus $r = 0.42$) in predicting wins from one season to the next.

4.2. WAR per Draft Position

The Jimmy Johnson draft chart [20] is widely cited in the NFL media and was based on the monetary cost of acquiring each draft pick. With PFF WAR at our disposal, we're able to look at history and determine how much, on average, should be expected from a player on his rookie contract by virtue of being drafted at a certain position in the NFL draft. Using the sum of a player's WAR for the first four years for each player drafted (regardless of whether they were playing for the team that drafted them) as a proxy for a player's rookie-contract WAR, we fit a LOESS [5] curve between draft position and the WAR and standard deviation in WAR (Figure 4). Notice, as one can expect, the payoff for picking a quarterback is, on average, higher than picking a non-quarterback. The variance is higher as well, though, and many quarterbacks taken in the draft are below replacement level as a result.

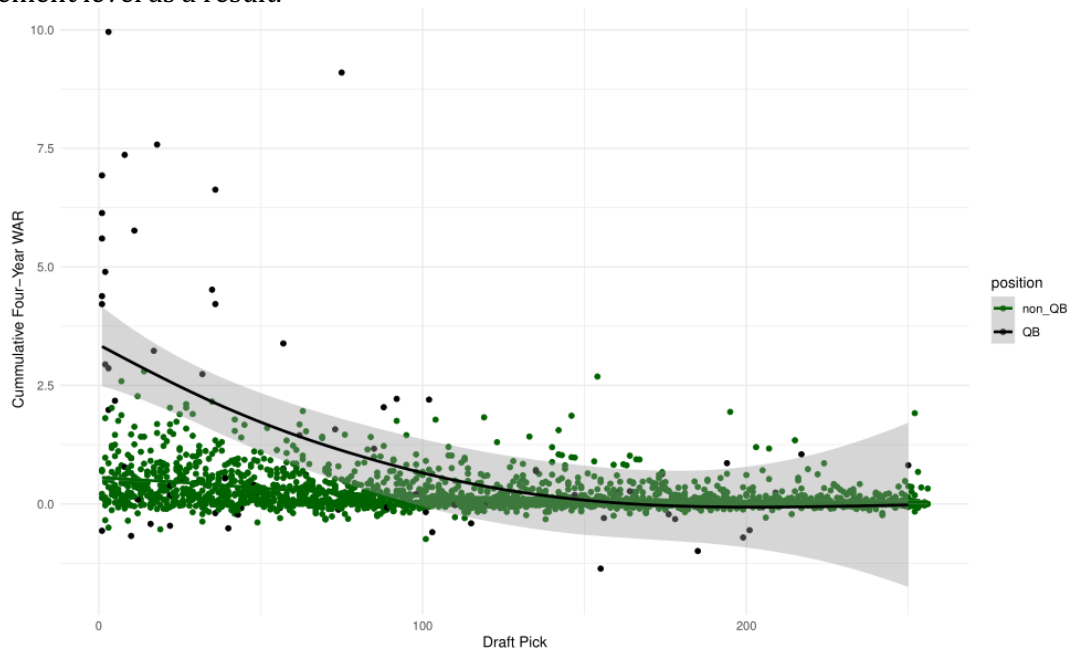


Figure 4: Early-career WAR as a function of draft position during the PFF era (2006-pres.), excluding players drafted in 2016 and beyond.

Since different positions offer different utility to different teams, the analyses of draft trades is a far more nuanced one that “the value of picks a , b , and c minus the values of picks x , y , and z ”. Take, for example, the 2018 pre-draft trade between the New York Jets and the Indianapolis Colts, which



netted the Jets the third-overall pick in exchange for the sixth-overall pick, the 37th-overall pick, the 49th-overall pick and a 2019 second-round pick (which turned into the 34th overall pick in the 2019 NFL draft). Assuming that the Jets *were open to taking* any position with the third pick, and you'd be left with the conclusion that, assuming a conservative [16] 25 percent discount rate on future picks, the Colts won the trade more than 75 percent of the time for an average almost one-and-a-half WAR during the course of the player's rookie deals (Figure 5). However, if we assume that the third-overall pick is a quarterback (which it was - the Jets picked Sam Darnold of USC) the Jets actually win the trade more than 65 percent of the time and generate roughly one-and-a-half WAR in the process. This exercise, which is along the same lines as the NFL's own Michael Lopez's research [14], shows that draft picks can have different utility for different teams. The Colts, possessing a franchise quarterback (at least at the time) in Andrew Luck, needed a volume of picks to re-stock their roster at other, less valuable positions. The Jets needed a quarterback and all the value that that brings and paid a substantial price for it with respect to the Colts. Both teams "won" that trade.

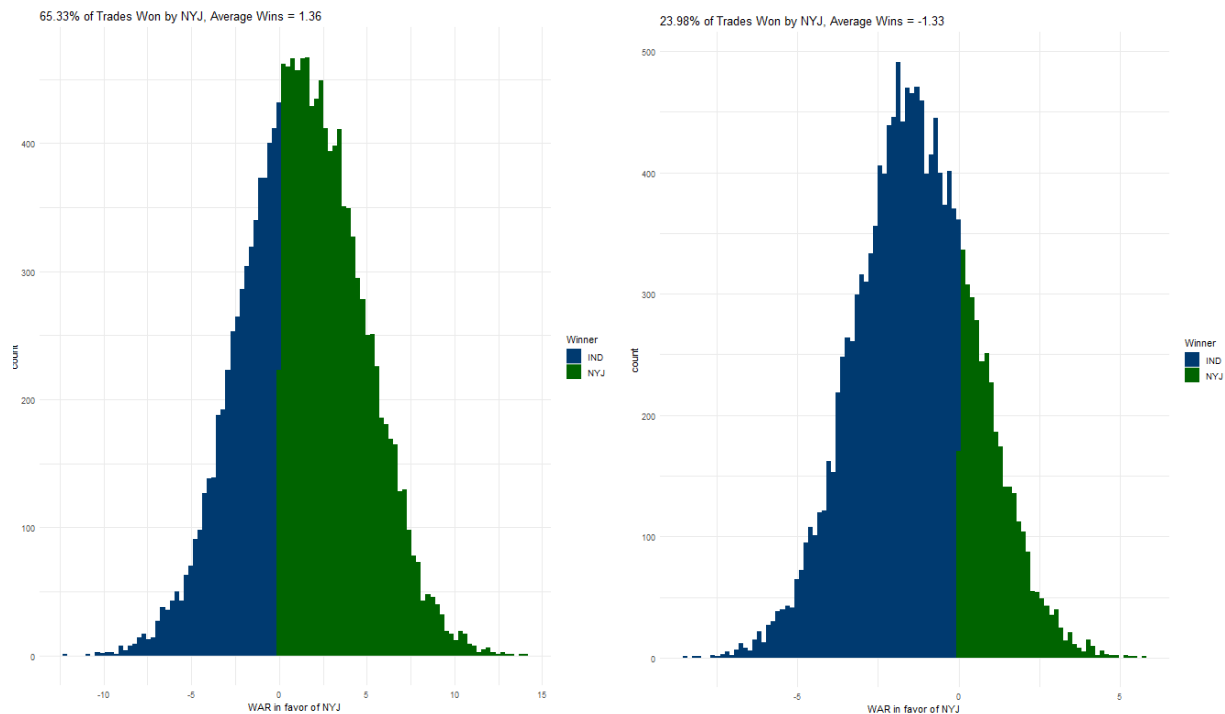


Figure 5: Outcome distributions for the 2018 Jets-Colts trade when assuming the Jets take a quarterback with the leading pick (left) and assuming they take a player from any position (right).

One can also look at expected draft position and determine which team did the best with the picks they used during the PFF era (Figure 6). Only drafts from 2006 to 2015 were used since the full four years for players drafted from 2016 onward have not been completed.

While there are some biases that skew this analysis (for example, a team that trades away or accumulates picks frequently), over this long a timeframe the results coincide with results on the field over the past decade and a half, with the Seattle Seahawks gaining the most WAR above



expectation and the Cleveland Browns the least. Even after quarterbacks are taken out of the analysis, the Seahawks are the top team and the Browns the bottom, but teams like Minnesota, Green Bay and New Orleans move ahead of teams that drafted quarterbacks who've had success (e.g. Atlanta).

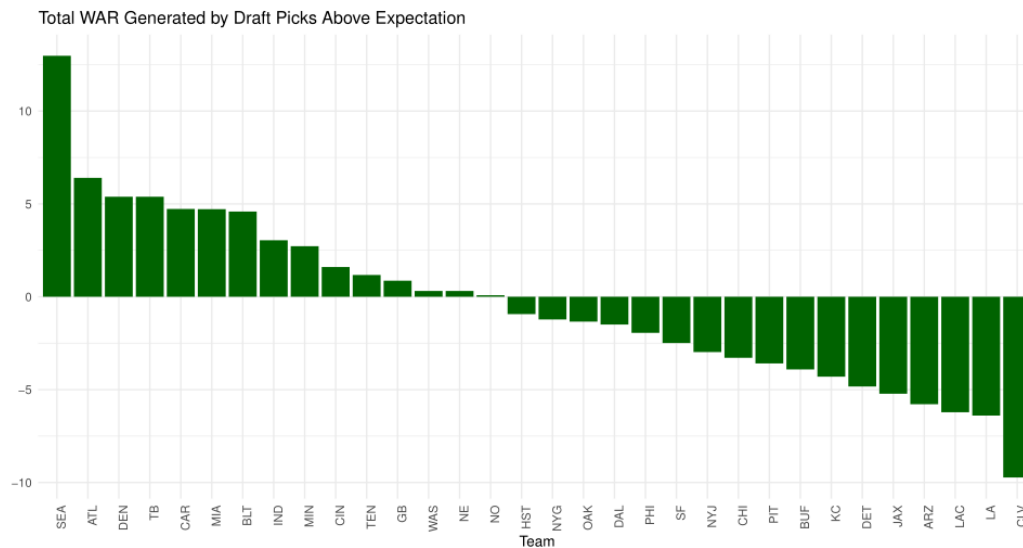


Figure 6: Early-career WAR accumulated by team over what is expected at each draft position. Quarterbacks are included in this analysis.

5. Discussion

In this paper we derived an estimate for every player's WAR at the NFL level using PFF data dating back to 2006. We found that, as expected, quarterbacks contribute the most to their team's success, and this contribution is more stable season-to-season than his other statistics. We also found that players further away from the football are, on average, more valuable to a team than players closer to the ball, but that value comes at the cost of predictability from year to year on the defensive side of the ball. We found that good tackles add more value than good guards and good centers, but the existence of dominant players in the latter groups make the mean values similar. The same holds true on the defensive line, where edge players are generally more valuable, but players like J.J. Watt and Aaron Donald have skewed the data in favor of interior players with their outlier play. It's very hard for a tight end or a running back to add the value of a high-level wide receiver, but equally difficult to produce values as far beneath replacement as the worst player at the position as well.

Season	Player	Position	Team	WAR
2011	Drew Brees	QB	NO	5.54
2017	Tom Brady	QB	NE	5.48
2006	Peyton Manning	QB	IND	5.17
2011	Aaron Rodgers	QB	GB	4.75
2006	Drew Brees	QB	NO	4.64



Table 2: Who has been the most valuable player in the PFF era? The top five players in terms of PFF WAR from 2006-present.

Team strength can be well captured by summing the collective WAR of its players, and this number is more stable season to season than actual wins or those implied by things like Pythagorean win percentages [11]. Teams that consistently overperform the sum of their PFF WAR values are those with coaches that are considered among the league's best, while teams that consistently underperform said numbers have been franchises with more than their fair share of dysfunction during the PFF era.

Player value decreases monotonically with draft position, as expected, but decreases differently for the quarterback position than it does for all other positions, rendering the historical Jimmy Johnson draft pick trade chart less meaningful than it was back when the passing game was less important. We demonstrate how the differential utility of a franchise quarterback can create "win-win" situations for trade partners on draft day, using the 2018 Jets-Colts trade as an example. Thus, while the seminal conclusions of Massey and Thaler [16] still hold when considering all positions, trading up for a quarterback appears to be a drastic exception.

Future work includes using these valuations of players to model and recommend salaries for players after their rookie contract is complete. The NFL has a rookie wage scale [18] that limits the amount that players in their first four or five years can make, independent of position and tied to where they were selected. Differences in the value of each position therefore manifest themselves in a) where a player is initially drafted and, more drastically, b) the duration, size and percentage of guaranteed money doled out in deals subsequent to their rookie deal. We find that offensive tackles and defensive edge players are paid more than their WAR numbers suggest, while quarterbacks, receivers and cornerbacks are paid less, so a model for recommending player compensation will likely need to be regressed to the market using positional factors for the time being. Models currently in development in this area are showing promise and will be the subject of a future manuscript.

Season	Player	Position	Team	WAR (rank)
2006	LaDainian Tomlinson	RB	SD	0.34 (113 th)
2007	Tom Brady	QB	NE	4.50 (1 st)
2008	Peyton Manning	QB	IND	3.66 (1 st)
2009	Peyton Manning	QB	GB	4.10 (1 st)
2010	Tom Brady	QB	IND	2.61 (8 th)
2011	Aaron Rodgers	QB	GB	4.75 (2 nd)
2012	Adrian Peterson	RB	MIN	0.30 (137 th)
2013	Peyton Manning	QB	DEN	4.45 (1 st)
2014	Aaron Rodgers	QB	GB	3.87 (2 nd)
2015	Cam Newton	QB	CAR	2.88 (5 th)
2016	Matt Ryan	QB	ATL	3.52 (3 rd)
2017	Tom Brady	QB	NE	5.38 (1 st)
2018	Patrick Mahomes	QB	KC	4.39 (1 st)



Table 3: The NFL MVPs during the PFF era, along with their PFF WAR values and where that WAR value ranked league wide.

Player valuation remains one of the most challenging problems in sports analytics, especially football. From the non-stationary nature of the sport, to the fact that different positions mean different things to different teams and schemes, the problem of assigning a value to each player in the National Football League will likely never be fully solved. However, with the aid of PFF data, which takes into consideration much of the vast and ever-changing play-by-play context that has alluded traditional statistics, we're able to estimate a value for each player in the currency in which they are ultimately judged on the football field: wins. This should have wide-ranging implications throughout the football and analytics community and be a jumping-off point for future approaches in this space.

Position	Player	Season	Team	WAR
QB	Drew Brees	2011	NO	5.54
RB	LaDainian Tomlinson	2006	SD	0.34
WR	Antonio Brown	2015	PIT	1.36
TE	Rob Gronkowski	2011	NE	0.68
T	Tyron Smith	2015	DAL	0.37
G	Evan Mathis	2013	PHI	0.46
C	Jason Kelce	2017	PHI	0.46
DI	J.J. Watt	2013	HOU	0.57
ED	Calais Campbell	2018	JAX	0.42
LB	Bart Scott	2008	BAL	0.39
CB	Darrelle Revis	2009	NYJ	1.19
S	Jairus Byrd	2012	BUF	0.86

Table 4: Highest WAR values during the PFF era by position. *Estimates adjusted since previous submission due to model updates.*

References

- [1] Baseball Reference. "WAR Explained". https://www.baseballreference.com/about/war_explained.shtml. 2019.
- [2] Basketball Reference. "NBA Win Shares". <https://www.basketball-reference.com/about/ws.html>. 2019.
- [3] Burke, B. "Does Defense Win Championships?" <http://archive.advancedfootballanalytics.com/2008/01/does-defense-win-championships.html>. 2008.
- [4] Clement, S. "The Relationship Between Punting and Winning in the NFL." <https://www.fielddgulls.com/2018/5/7/17327282/punt-wins-above-average>. 2018.
- [5] Cleveland, W.S. "LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression". *The American Statistician*. 1981.



- [6] Corsica. "The Art of WAR." <http://www.corsica.hockey/blog/2017/05/20/the-art-of-war/>. 2017.
- [7] Eager, E., Chahrouri, G. "Edge vs. Interior: Which Pass-Rusher Reigns Supreme." <https://www.profootballfocus.com/news/pro-edge-vs-interior-which-pass-rusher-reigns-supreme>.
- [8] Eager, E., Chahrouri, G. "PFF Data Study: Coverage vs. Pass Rush." pff.com/news/pro-pff-data-study-coverage-vs-pass-rush. 2019.
- [9] ESPN. "Introducing a New Way to Rate College Football Players". espn.com/college-football-story/_/page/PSimpactrating0819/introducing-new-way-to-rate-college-football-players. 2019.
- [10] ESPN. "How is Total QBR Calculated? We Explain Our Quarterback Rating." espn.com/blog/Statsinfo/post/_/id/123701/how-is-total-qbr-calculated-we-explain-our-quarterback-rating. 2016.
- [11] Football Outsiders. "Football Outsiders Almanac". 2011.
- [12] Ho, T.K. "Random Decision Forests". "Proceedings of the 3rd International Conference on Document Analysis and Recognition". 1995.
- [13] Hockey-Graphs <https://hockey-graphs.com/category/war/>. 2019.
- [14] Lopez, M. "Rethinking Draft Curves" <https://statsbylopez-netify.com/post/rethinking-draft-curve/>. 2018.
- [15] Massey, K. "Statistical Models Applied to the Rating of Sports Teams". Bluefield College Honors Thesis. 1997.
- [16] Massey, C., Thaler, R.H. "The Loser's Curse: Decision Making & Market Efficiency in the National Football League Draft." *Management Science*. 2013.
- [17] McLean, R. A., Sanders, W.L., Stroup, W.W. "A Unified Approach to Mixed Linear Models" *The American Statistician*. 1991.
- [18] Over the Cap. <https://overthecap.com/>. 2019.
- [19] PFF. pff.com. 2019
- [20] Pro Football Reference. <https://www.pro-football-reference.com/>. 2019.
- [21] R Core Development Team. "R: A Language and Environment for Statistical Computing". 2019.
- [22] Rodgers, J. L., Nicewanter, W.A., "Thirteen Ways to Look at the Correlation Coefficient". *The American Statistician*. 1988.
- [23] Slowinsky, S. "What is WAR?" <https://library.fangraphs/misc/war/>. 2010.



[24] WAR on Ice. "The Road to WAR (For Hockey), Part 1: The Single-Number Dream." [Http://blog.war-on-ice.com/index.html?f=37.html](http://blog.war-on-ice.com/index.html?f=37.html). 2015.

[25] Yurko, R., Ventura, S., Horowitz, M. "nlWAR: A Reproducible Method for Offensive Player Evaluation." *Journal of Quantitative Analysis of Sports*. 2019.