Fall 2021

Trade Bait: Season 3

Ben Bagley
Western Washington University

Trade Bait, Season 3: A dissection of statistics in the NFL and NFL Media
By Ben Bagley, Advised by Derek Moscato
WWU Honors Capstone Project, Fall 2021

Episode 1
NFL QBR and the issue with stats in the media

Hello and welcome to the third season of the Trade Bait podcast! My name is Ben Bagley, and this third season is serving as my senior capstone project for the honors program at western Washington university.

I wanted my senior project to be a culmination of the education I have received both in and outside of the classroom. I am set to graduate with a double major in statistics and journalism, and it has been my professional interest to marry advanced analytics with multimedia sports reporting. Through college, this sports stats podcast has been my platform to do this. I've published two seasons talking stats about professional football, baseball and basketball, and I have learned quite a bit about the nuance and intricacies of talking sports stats to an audience. And now, I'm very excited to present to you Trade Bait, Season 3: A five-episode series dissecting the use of statistics in the NFL and NFL Media.

(intro music)

Hello again and welcome to episode 1 of the Trade Bait podcast, season 3. This season is specifically dedicated to dissecting the use of statistics in professional football, both by teams and by the media. As a fan and now a young professional statistician, I've always been enamored by analytics and NFL next gen stats and all the numbers that tell the story of the game. However, I've also felt increasingly dissatisfied by the use and calculations of some modern metrics. I think there's a missing link when discussing the value and skill level of certain positions in football, especially on offense. Even with the modern NFL using analytics more than ever before, our evaluation of wide receivers, running backs and even quarterbacks – the most important position of them all – fall flat. In this first episode we'll be taking a look at the quarterback position and breaking down his most fundamental statistic: NFL passer rating.

(transition music)

Passer rating in the NFL is a statistical attempt to assign a numerical value to a quarterback's performance. Based on his in-game stats, a passer rating tries to quite literally rate a quarterback to say how good he played. A high passer rating indicates a good game, and a low passer rating a bad one. But before we break it down and start discussing its pros and cons, we need to first understand the history of the metric.

Passer rating was first formally adopted by the NFL as the official metric for rating quarterbacks in 1973. Two years earlier, in 1971, NFL commissioner Pete Rozelle was frustrated with the lack

analysis that simple passing metrics like yards and completion percentage were providing, and asked the league's statistical committee for a better metric. This committee was headed by Don Smith, a public relations employee for the Pro Football Hall of Fame, Seymour Siwoff who was head of Elias Sports Bureau, and Don Weiss, an NFL public relations executive and head of game day operations at the Super Bowl.
Initially, the metric used four variables: percentage of completions per attempt, average yards gained per attempt, percentage of touchdown passes per attempt, and percentage of interceptions per attempt. Notice that this metric wasn't plugging in raw stats such as yards and touchdowns, but rather examining results per attempt.

Because the calculation was so simple, many other statisticians attempted to improve upon it over the next decades. Many critics of the metric proposed two main issues: for one, it didn't consider the running ability of quarterbacks, which has become exceptionally common in the modern NFL. Secondly, it didn't matter if the team won or lost, but the accumulation of more stats just naturally led to a higher rating.

Inspired by these pitfalls, David Berri, a sports economist from Utah State proposed the QB Score in 2007. He analyzed data to determine which aspects of a game contributed most to a team's scoring or surrendering points (and therefore winning) and designed the QB Score to capture how well a QB contributed to a win.
This was calculated as Yards minus the amount of offensive plays times three, minus the amount of QB turnovers by 30. His idea was that running more plays and generating more yards translated to more points and wins, and less plays paired with more turnovers translated to fewer points and wins. This metric was all about efficiency. It didn't reward quarterbacks for piling up the stats, but rather for playing cleanly, efficiently and with intelligence.

Passer rating, however, continued to be used and continued to be tweaked. Paul von Dohlen, a mathematics professor at William Paterson University, tweaked the metric in 2011 by replacing completion percentage with yards per attempt. This was in an effort to reward quarterbacks for pushing the ball downfield rather than taking short, simple passes. Following this line of thinking, Ahmed Cheema, a writer for The Spax took it a step further and proposed the Adjusted Passer Rating in 2019.

The main difference is that APR uses net yards per pass attempt instead of passing yards per attempt. Net yards per attempt factors in the quarterback being sacked: (passing yards – sack yards) / (passing attempts + sacks). Completion percentage was also removed from this model in order to not reward quarterbacks for passes that are completed for zero or negative yards.

The metric has been far from perfected, but as of 2021 Passer Rating is the official NFL statistic for evaluating quarterback performance. The worst rating a quarterback can get is 0.0 and the best, called a "perfect passer rating" is 158.3. It's almost infamous that the metric is on a 0-158.3 scale, and if you're anything like me, that scale has always been incredibly confusing. So let's break it down, here's how it is calculated today.

(transition music)

Alright bear with me as we get nitty gritty with some numbers:

First, you take five raw statistics: Completions, Yards, Touchdowns, Interceptions, and attempts. You then translate that into four variables, completions per attempt, yards per attempt, touchdowns per attempt and interceptions per attempt. Each of these variables are then scaled to have equal weight so they are each equally considered in the final calculation. Keep in mind that a typical number of touchdowns could be about 3, the number of completions could be 30 and the average number of yards could be 300, so scaling is needed to balance out these raw numbers. Here is how the scaling is done:

Variable A
Touchdowns per attempt is multiplied by 20. If a player throws 3 touchdowns in 30 attempts, that .1 touchdowns per attempt becomes a variable value of 2.0 with the scaling.

Variable B
Completions per attempt minus 0.3 and multiply that by 5. If a player completes 20 of 30 passes, the variable calculation would be: 20 out of 30 minus 0.3 times 5. In this example the variable would come out to 1.83, which we see is on a similar scale to that 2.5 in the touchdowns per attempt variable.

Variable C
Yards per attempt minus 3 and multiply that by 0.25. If a player threw for 300 yards on 30 attempts, the variable calculation would be: 300 divided by 30, minus three, times 0.25. In this example the variable comes out to 1.75, which we see falls on that same scale.

Before we get to the final variable, there is one other point to consider in this scaling calculation. We see that the numbers that come out hover in the 1.5 to 2.5 range. This metric actually scales on a 0 to 2.375 scale. This means that the maximum value for any variable is set to be 2.375. If the variable value comes out to be greater than that, say, 2.5, it is simply set as 2.375. The 0 is a hard floor, too: if any variable comes out to be negative, it will simply be set to 0. So the maximum value for any of these variables is 2.375 and the minimum is 0.

Now, to Variable D
Interceptions are a negative thing for a quarterback, and the most ideal number is to have 0 interceptions. So, to figure out the variable, we take that 2.375 and subtract our scale: in this case it is interceptions per attempt times 25. Say a player threw 1 interception in 30 attempts, the variable value is 2.375 minus (1 out of 30 times 25). This would come out to 1.54, which fits the scale. Note that if a player throws zero interceptions the variable come out to be that max value of 2.375, because it would be 2.375 minus (0 out of 30 times 25) which is 2.375 – 0 which is 2.375.

Now we're on to the final step of putting it all together.

First, we add up all the variables and divide by six. Now, I did my research and that 6 appears to just be an arbitrary value, so just accept it with me. Take all the variables, divide by six, and multiply it by 100. The final result you get is your passer rating! For our example, for a quarterback that completed 20 of 30 passes for 3 touchdowns, 1 interception and 300 yards, their passer rating would be 118.75. The league average as of 2020 was 93.6, and so we see that our quarterback's above-average performance translated to an above-average passer rating.

How does one get a perfect passer rating then, you ask? A player would have to have the max value of 2.375 for each variable, as 2.375 divided by 6 multiplied by 100 gives us the infamous 158.3 passer rating. To get a max value for each variable, a quarterback would need to have AT LEAST the following: a completion percentage of 77.5 percent, 12.5 yards per attempt, one touchdown for every 8.42 attempts, and zero interceptions.

For the quarterback in our example to have had a perfect passer rating his 30 attempts, he would have had to completed 24 passes for 375 yards with four touchdowns and no interceptions. Now that we know how it is calculated, we can begin to nitpick at its effectiveness just like our statistical forefathers.

(transition music)

So by going through this calculation we can see: the true definition of a "perfect" game for a quarterback in terms of the official NFL statistics is simply having a completion percentage of 77.5 percent, 12.5 yards per attempt, one touchdown for every 8.42 attempts, and zero interceptions. It doesn't matter if you won or lost, it does not take into account the quarterback's running ability, and it doesn't take into account any kind of situational passing, where a 5 yard completion on 3rd and 3 would be much more valuable than a 5 yard completion on 3rd and 10.

Following that order, here's my first issue with passer rating: Running quarterbacks are not rewarded. Yes, I realize that it is very specifically called *passer* rating, but the fact that it is the official NFL stat for measuring quarterbacks makes this problematic. In Week 2 of this NFL season, Lamar Jackson – a terrific running quarterback who set the single-season qb rushing record in 2019, defeated Patrick Mahomes and the Kansas City Cheifs by a final score of 36-35. Jackson threw two interceptions in the first quarter and finished the game with the final line 18 of 26 for 239 yards, 1 touchdown, 2 interceptions and a mediocre 78.8 passer rating.

However, he also rushed 16 times for 107 yards and two fourth-quarter touchdowns, and came back from an 11-point 4th quarter deficit to win the game. Based on passer rating, he had a subpar game, but his running ability from the quarterback position is what won them the game.

In 2020, Arizona Cardinals quarterback Kyler Murray finished 18th in passer rating but led all quarterbacks with 11 rushing touchdowns and was second with 819 rushing yards (17th among all players). Despite the all-around production, where he also had 4,000 passing yards and 26

passing touchdowns, his 12 interceptions on the year meant that he finished with a worse Passer Rating than Dolphins QB Ryan Fitzpatrick and Jaguars QB Gardner Minshew, each of whom threw for less than 2,300 yards, less than 20 touchdowns, and rushed for about 150 yards on the year with three rushing touchdowns combined. According to the Passer Rating formula they were "better" quarterbacks than Kyler Murray, even though the stats show they clearly were not.

I don't think that rushing statistics should be heavily added into the passer rating formula, but we have to acknowledge its value. There are great quarterbacks that hardly run like Tom Brady and Aaron Rodgers, but quarterbacks that can do it both should be rewarded. It's also difficult when we look at this as members of the media, because there *isn't* a statistic that adds passer rating to runner rating. When comparing somebody who runs so often like Lamar Jackson or Kyler Murray to somebody who strictly throws like Brady, Aaron Rodgers or Drew Brees, it's honestly comparing apples to oranges. We will look at passer rating and say the throw-heavy players are better, but that simply is not the case.

Here's is my second issue with passer rating, and it was best stated by Pro Football Focus writer Nathan Janke: passer rating rewards safe quarterbacks who may not be actually helping the team. Janke explained it as this:

If Quarterback A completes three passes for three yards in a row, they would have a passer rating of 79.17. Meanwhile, if Quarterback B throws three straight passes, with the first two falling incomplete and the third being caught for a 30-yard gain, they would have a passer rating of 71.53. While Quarterback A has a higher completion rate and is more efficient than Quarterback B, they have failed to gain 10 yards on the drive and reward their team with a first down. On the other hand, while Quarterback B wasn't necessarily efficient in their passing, they were able to make a completion of 30 yards on their third pass, thus giving their offense a first down and a significantly improved field position (likely on their opponent's half of the field).

On the flip side, a quarterback who is highly accurate but doesn't go deep and rather throws a lot of short, efficient passes like slants, crossing patters or screens but is highly effective would actually be punished by that minimum 12.5 yards per attempt you need for a perfect passer rating. In a game in 2019 against the Indianapolis colts, New Orleans Saints quarterback Drew Brees completed 29 of 30 passes for 307 yards, 4 touchdowns and 0 interceptions. His 96.7 completion percentage in the game set the NFL single-game record in a dominant 34-7 victory. However, he did not finish with a perfect passer rating because he only averaged 10.2 yards per attempt. So for a quarterback who quite literally came as close to a completely perfect NFL game as any other player in history, he did not have a perfect passer rating. But Eli Manning, who in 2009 for the New York Giants completed 8 of 10 passes for 173 yards and two touchdowns, did. That's right, he completed 8 passes for less than 200 yards but, according to the math and earned a perfect passer rating, but the all-time most accurate game in NFL history did not. That in and of itself is indicative of an err in the calculation.

Following this line of thinking, passer rating actually rewards modern NFL quarterbacks because the game has moved to such a pass-centric state of offense. Because of this, Matt Schaub, who has been a backup most of his career with about three good seasons in the early 2010's actually has a higher career passer rating than 90s quarterbacks Dan Marino and John Elway, each of whom are in the hall of fame and have a combined 19 pro bowl selections and 2 super bowls. If the calculation of this metric can be used to argue that Matt Schaub had a better career than Dan Marino and John Elway, then that's more proof there's an err in the calculation.

Finally, the biggest thing that passer rating does not quantify is situational football.
My idea of a perfect passer rating, when it includes the word perfect, means that every throw was completed for a touchdown. Not only that, but it would mean that every throw was a 99 yard touchdown to maximize yardage and a score on every single throw. But we've already explained that the metric's idea of perfection just means that a quarterback's yards, completions touchdowns and interceptions per attempt eclipsed a predetermined number. But now, if we have two different quarterbacks who each have a perfect passer rating, how could we possibly distinguish between the two? How do we discern which one was better if they were both perfect? Would you say that Eli Manning's 8-10 with 170 yards and 2 was better than Drew Brees' 29 of 30 for 300 yards and 2 scores because of passer rating? The answer, to me, is that we have to look situationally at each decision a quarterback made.

On every passing attempt, a quarterback usually has 2-5 pass catchers at his disposal. Each is running their own route against a unique defense and a quarterback has a decision on where he wants to put the ball. Say he spots his tight end over the middle and hits him for a 15-yard gain. But say on that same play he also had his wide receiver wide open, uncovered 45 yards down the sideline available for a touchdown – but he didn't see him, so didn't throw it. He still completed a pass for enough yards to be "perfect" in terms of passer rating, but his decision to hit the tight end eliminated the possibility of a touchdown on the play. In my opinion, his rating should actually take a hit because he didn't make the perfect decision. He made a correct decision, but not the most correct decision.

Imagine now a third down scenario, or rather two third down scenarios. In both situations it is third and three and the quarterback completes a 5 yard pass for the first down. However, in situation A, it's the first quarter of a 0-0 game. In situation B, there is 30 seconds left in the 4th with your team down by 3 and you just converted a third down that puts you in field goal range to tie the game. Imagine a quarterback who completes that pass in the first but not in the 4th compared to a quarterback who misses it in the first but completes it in the fourth. It adds up the same in the final stats, but situationally the 4th quarter 3rd down was much more important to the game but the qb who completed it doesn't get rewarded for his clutch play.

This is where passer rating fails in an extreme way. There is no way to discern between situations, right or wrong decisions with the football, or any of that stuff. It doesn't even take into account the amount of yards the ball traveled in the air, so a little dump pass to a running back who then runs for 50 yards after catching it is treated the same as a 50 yard dime to a receiver in stride down the sideline.

What I'm trying to say is that as much as I love statistics they don't tell the whole story when evaluating quarterbacks. We need to be able to incorporate the situation, the decision and the they made, and then other factors like how open the receiver was, what play the defense was running, how well the offensive line protected him, all kinds of things. And so when we use such a one-dimensional metric in the media, the conversation just lacks depth. Passer rating is an interesting metric that does tell a story, but it's far from a perfect evaluator. I think it can certainly be used NFL discourse and the media, but we have to be sure we're telling the full story behind the stats. Because if we don't, we run the risk of hailing Matt Schaub as a better quarterback then Dan Marino. The media plays such a big role in how fans and honestly front offices perceive quarterbacks, regardless of stats, so if we're using metrics in our discourse that are not well rounded its almost irresponsible, and I think we have the tools to be better than that.

The last gripe that I have with passer rating is that it's on a scale of 0 to 158.3. Nobody thinks on a scale of 0 to 158.3, we think on scales of 0-10. 0-5. When you leave a Yelp review you don't say "Oh I thought this restaurant was a 84.6 out of 158.3", you say its 4/5 and we as Americans understand that fraction. It's literally so easy to put passer rating on a 0-100 scale, in the calculation when you divide the sum of the variables by 6, if you instead divide by 9.5 then a perfect passer rating would be 100, and if I told you "his passer rating was 80" you understand where it falls on that scale. I genuinely do not know why we as the NFL media haven't taken care of this yet, but that's definitely something that is just confusing to absolutely everybody and has to be changed.

(transition music)

Aside from my gripes with Passer Rating, it is still a useful statistic for a baseline evaluation of NFL quarterbacks. I do, however, want to present some other metrics currently on the market and used by NFL Media members that I think do a good job fixing some of these issues.

The first is ESPN's QBR, or Quarterback Rating. This confused me for a long time, but QBR and passer rating are different metrics. This metric, released in 2011, is a measure of efficiency that looks at a quarterback's contribution to winning. Not only does it use the basic NFL statistics, but it also assigns credit to the quarterback and other teammates to determine how much of the play the QB was responsible for. It looks at performance on every play and, by taking into account the down and distance, quarter, score, time remaining and other factors, and estimates expected points per play. This weights clutch scenarios like a third down in the red zone or a key interception, and doesn't care as much about plays like a hail mary interception at the end of the half or garbage time stats when a team is already up 40. This is a statistical attempt to quantify that situational football that I was talking about.

They also divide credit between players. So, for example, a quarterback who threw a 50 yard pass rather than a screen to the running back that went for 50 yards would get more credit for the play. The expected points added per play is divided up between the quarterback, receiver

and offensive line to account for this. All of this data is charted and tracked by a team of ESPN Stats and Information analysts, and is far outside the realm of possibility for statisticians like me right now. Many of these metrics, including NFL next gen stats, also utilize player tracking data and many more advanced (and expensive) means of gathering data.

ESPN QBR is then put on a 0-100 scale, where 50 is an average quarterback and 75 is pro-bowl level. I like this metric a lot, for obvious reasons. It addresses a number of issues that I had with Passer Rating, and it even puts it on a very readable scale. One issue, however, is that is takes into account statistics that are not available to the average fan, like expected points added per play. ESPN themselves have never actually released the formula for QBR, but rumor has it that it's made up of over 10,000 lines of code. In theory it is a very creative way to get around the issues Passer Rating has, but it also could be seen as a random soup of statistical processes that just further complicate everything. But, even if it needs some perfecting, it's a great answer to many of the questions and issues I had with Passer Rating.

The other metric that I wanted to highlight is from Pro Football Focus. Pro Football Focus (PFF) is a private football website that conducts in-depth analysis on NFL and assigned a grade that indicates the performance of a player.

According to the PFF website, the group's algorithm analyzes every play for each individual player and measures the impact that said player has while on the field. A player's impact is then given a grade between −2 to +2 in 0.5 increments. Each position has a scale with a unique algorithm and rules. The scale also takes into account game context, so a strong play in the fourth quarter of a close matchup will be graded higher than one in the 2nd quarter of a blowout game.

A 0-player grade on any given play represents any position player performing at an expected level and in a manner that neither positively nor negatively impacts their team. An example of this is a running back taking a carry through the correct hole and picking up 3 – 4 yards on first and ten. Meanwhile, a +2 represents an incredible performance on a play that shifts the dynamic of a game in favor of the player's team. Brandon Graham's, a defensive end for the Philadelphia Eagles, strip sack on Tom Brady in Super Bowl LII would have easily been a +2 rating.[16] On the other hand a −2 is a play that catastrophically hinders a team's chance of winning, such as a quarterback throwing a pick-six in the fourth quarter of a close game. Ambiguous plays where the outcome is unclear on how a player impacted their team are typically given a 0.

The sum of these plus-minuses are then converted on a 0–100 scale and produce a grade for a single game. However, a player's season grade is not the average of the 16 grades a player receives each game. Instead, PFF credits a player's entire body of work and longevity throughout the season. It is, therefore, possible for a player to have a higher season grade than any individual grade that a player received in any game he played in.

This metric is also fun in that it incorporates situational football, but it doesn't focus as much on raw NFL statistics and falls into that category of being one step removed from NFL fans. It's a very arbitrary rating from a private football website, so it can be tough to justify a player's score. Also, PFF operates behind a paywall so fans need to pay to have access to their player grades. It's great work by clever statisticians, but difficult to use in media settings. But, just like QBR, it provides an interesting alternative to Passer Rating that accounts for some of its pitfalls.

(transition music)

That's it for the first episode of the third season of Trade Bait! I hope you learned something about the NFL's official statistic for quantifying quarterback success, Passer Rating, and I hope you too have some questions about how we can better analyze on-field success. QBR and Pro Football Focus player grades are exactly that, the result of curious people trying to improve upon what we have. Talking statistics in NFL media is an ever-evolving artform, and it remains important that we understand the stats we're using and can contextualize what they tell us. Be sure to think critically about the stats that are being presented to you, and always try to understand what story a specific metric is talking about. The more intentional we are in the selection of our statistics, the more accurate we can be in our discussions. Next episode we will talk about a different way of evaluating quarterbacks by looking through the lens of financial value and I'll introduce you to my new metric, the Bagley Value Rating. Thank you for tuning in to the first episode of my senior capstone project, and I'll see you soon for episode 2.

Episode 2
BVR Score

Hello again and welcome to Trade Bait Season 3, Episode 2. In the last episode we looked at NFL Passer Rating and we talked about the overall evaluation of quarterbacks. With that stat, and we also mentioned other metrics like ESPN's QBR and Pro Football Focus player grades that evaluate quarterbacks in much more advanced ways, and we could really go on and on with different stats that try and quantify a quarterback's performance in different ways. There are tons of them out there, and a lot of creative statisticians have produced very interesting content over the year.

In April of 2020, however, I was struck with an idea for a different approach for measuring a quarterback's value. What sparked that? Patrick Mahomes. On April 30th, 2020, the Super Bowl winning, MVP quarterback for the Chiefs was rewarded for his phenomenal start to his career by signing a 10-year, four hundred and seventy-seven-million-dollar contract that could inflate to over 500 million dollars with added bonuses. It was the largest contract in American sports history, and he also became the first athlete to half a half-billion-dollar contract.

What struck me about that deal was not the size of it. Quarterback salaries have been ballooning since the 2010s and it was only a matter of time before some really good young player got a contract that was worth that much. But what struck me was the fact that the Chiefs would be paying him that much money while still needing to operate under the NFL salary cap. And not only that, but he signed that deal in 2020, when revenue across all professional sports leagues were down due to the COVID-19 pandemic.

So let's quickly talk about the salary cap. The NFL has a hard salary cap, which means that there is a set amount of money that teams can spend on their players. The NFL is a league with a hard salary cap. The cap marginally increases or decreases every year based on league revenue, TV deals and the collective bargaining agreement between the NFL and the NFL players association, but it is how much money every team has to pay the salaries of their players. The NBA has a soft salary cap, which means there are certain situations in which a team can spend more on their roster than the cap and pay extra in taxes. The MLB has no salary cap, so teams can spend as much as they want on players, which usually results in teams in bigger markets having better teams. There's a reason why the Yankees and Dodgers and Red Sox are good every year, because New York and LA and Boston just have more money than Oakland and Seattle. But in the NFL, every team has the same amount of money to spend on players, no matter what.

When the latest CBA – collective bargaining agreement – was signed in 2011, the NFL salary cap was set to 120 million, 375 thousand dollars. The highest paid player in the league was New York Jets quarterback Mark Sanchez, who's salary was just over 17 million. Since then, the salary cap grew by about 7 percent every year as the league continued to generate more revenue. By 2015 the cap was up to 143 million, and in 2018 it hit 177 million and the highest paid player was San Francisco 49ers QB Jimmy Garoppolo at $37 million. In 2020, the salary cap

hit an all-time high at just over 198 million dollars. With the cap space growing by nearly 80 million dollars in just 10 years, it's clear to see why quarterback salaries have continued to grow as well. It's also worth noting that, due to the loss of revenue in 2020 because of COVID, the 2021 salary cap actually decreased by 15 million dollars, the first time the salary cap has ever decreased since it was first introduced in 1994.

So under the umbrella of the cap, which *usually* grows from year to year barring a global pandemic, teams have the freedom pay players any salary under the cap value. However, an NFL roster consists of 53 players, so teams must be calculated in dividing up the funds in order to have the most complete, well-rounded roster possible. Players of a higher skill level (which can be determined by advanced metrics) typically demand higher salaries, so teams are in a tricky position when deciding how much to offer a player. If a team pays a player too much, then they won't have enough money to round out other positions. If a team offers a player too little, the player is unlikely to sign with the team and will go to a different team.

Quarterbacks are widely considered the most influential position to the success of an NFL team. As historical trends have shown, teams with good quarterbacks enjoy success, and teams with bad quarterbacks suffer failures. Therefore, since quarterbacks are the most important, they are the highest paid position. Teams want the best QB possible, but the better the player, the more money the team must pay them. With this in mind, the obvious conundrum materializes: teams desire the best quarterback possible, but also want to avoid paying too much in order to have enough salary cap room left to fill out the best overall roster possible.

This is the line of thinking I went down last winter, and I think you can see where we're going with this. I asked the question "which quarterback will give me the most production on the lowest salary?" And vice versa, which quarterbacks are giving me the least amount of production based on a high salary? This is the idea of value in this project. The most financially valuable quarterbacks are ones who can put up statistics that far exceed their salary. The least financially valuable are players whose statistics are far worse than their salary. So, which quarterbacks are the most valuable? The best quarterbacks are the ones paid the most, but do they give teams the most bang for their buck? Which quarterbacks are the least valuable and most overpaid, and are actually ripping their teams off most? Young quarterbacks can have extremely high values in their first years playing on a smaller rookie contract, but veterans playing at an elite level can still provide value with a large contract. So, who's worth it?

This is where the Bagley Value Rating (BVR, pronounced "beaver") Score comes into play. This metric takes a quarterback's stats and salary and, based on other quarterback's numbers across the league, manufactures a "score", which arbitrarily evaluates that player's financial value. By identifying the most and least valuable players, we can add a deeper level of evaluation to NFL discourse. While other advanced metrics work to figure out which QB is best overall, the BVR Score adds a new dimension by figuring out which quarterback is the most valuable, giving the team the most (and least) bang for their buck. Allow me to present to you, the Bagley Value Rating

(transition music)

The BVR Score uses both in-game statistics and yearly salary to determine the results. It is important to know that this metric is not static, as salaries change year to year. For this podcast I updated the algorithm with the statistics and salaries of every starting QB's game through Week 10 of the 2021 NFL season. From there, I could generate a raw score for every player in every game and from there I could create a scatterplot with salaries on one axis and scores on the other. After that I was able to run a linear regression to find a line of best fit through the data, which gives an estimate of how all the raw scores across the league are related. From there, I was able to determine if a player was over or underpaid based on their stats and salary compared to the rest of the league.

The first step in the project was determining which statistics to use in the calculation of the raw scores. After our discussion about passer rating, I decided to include seven variables: Net Yards per Attempt, which considers sacks and sack yards along with passing yards, Interception Rate, Touchdown Rate, Yards per Completion, Yards per Rush, Touchdowns per rush and, of course, salary. Let's quickly break down these variables. Interception and touchdown rate measures the percentage of throws that go for scores or picks to account for quarterbacks that throw more or less than others. Rushing statistics are included for quarterbacks who tend to run, but they aren't added in a way that punishes less mobile quarterbacks. For example, having no rushing yards will not hurt a player and kneels at the end of games aren't included, but having a rushing touchdown or some yards would boost your score a bit. Also of note, the salary included is a player's salary for the given year – called a cap hit – rather than their entire contract. So the salary variable for Patrick Mahomes isn't 504 million dollars, but rather what he is getting paid in 2021 – about 7.5 million dollars since it's the final year of his rookie deal.

(transition music)

To calculate a player's raw score, we scale and add up those variables. We take net yards per attempt + touchdown rate + yards per completion + yards per rush + touchdowns per rush – interception rate. By scaling, and maybe a better word is weighting the variables, I mean that, for example, touchdown rate and interception rate is multiplied by 3000 to turn a number like 0.025 into 75. Net yards per attempt is multiplied by 10 to turn a number like 5 to 50 – so we just want each of these variables to be equally important in the final calculation. Salary is also divided by a million to turn 30 million into 30.

Every starting quarterback for every game will be plugged into this algorithm to generate a raw score. The x axis, the horizontal one, is salaries and the y-axis is are the raw BVR scores. When we create our initial scatterplot, raw BVR scores are pretty equal across the league, it's a very flat linear relationship. We can see that, depending on salary, quarterbacks largely put up similar numbers. As our x-value, salaries, increase, on-field stats stay largely the same. This makes sense when we realize that somebody like Aaron Rodgers is making 27 million dollars, Dak Prescott is making 17 million dollars, and Patrick Mahomes is making 7 million. In any given game you could expect each of those guys to put up equally good numbers, so our scatterplot

makes sense. This also makes sense the other way, when somebody like Ben Roethlisberger is making 25 million dollars, Baker Mayfield is making 10 million, and Trevor Lawrence is making 6 million. On any given day this season, you could expect any of those guys to put up equally mediocre numbers. So initially we can see that typically a quarterback's salary does not influence their performance.

A quarterback's salary does, however, influence their value. This is where we introduce weighting. I created another scatterplot, but this time, instead of salaries on the x axis and raw BVR scores on the y axis, which gave us that very flat linear relationship, I put salaries on the x and BVR scores **times** salary on the y. This introduces weighting into the scores. Players with higher salaries now produced higher scores and players with lower salaries produced lower scores. They could have identical stats, but the multiplying by a higher salary inflates the scores of those highly paid players. From here I ran a linear regression that generates a line of best fit through the data, which estimates the league average BVR score for every possible salary, with the weights of the salary included. This line of best fit was positive and increasing, so higher salaried quarterbacks were now expected to have higher scores.

I then took that line of best fit, which was sloping upwards in the positive direction, and put it back on our initial, flat, linear scatterplot. Now, I had a scatterplot of scores from around the league that didn't depend on salary, and a line going through it that did consider salary. This is what I called my line of value. Now, because we know qb stats are largely similar across the league despite salary, we have a weighted line that allows for lesser paid quarterbacks to be more valuable than higher paid quarterbacks when putting up identical numbers.

Each point in this scatter plot will have a residual based on that line of value. A residual is how far above or below the line of value a point is. A positive residual means the actual value is greater than the expected value, and a negative residual is the opposite.

To find the true BVR Score for a player, we calculate the mean (average) of their **standardized** residuals. While residuals are how far above or below a point falls from its expected value, standardized residuals normalize the data and instead quantifies how large the residuals are in standard deviation units. A very large, positive standardized residual means that the data lies well above the expected value, and large but negative means the data lies well below. A value of 0 means that the actual value lies right on the mean. In our context, that would mean that a player is performing at exactly his salary level. Positive values mean over performing, and negative values mean underperforming. The reason that we standardize it is so we can discuss players of all values on the same normalized scale.

And, that's it! A player's final BVR score is the average of their standardized residuals, multiplied by 10 for readability. There is no scale of the metric, no "perfect score". A positive BVR score means that a player is out performing their contract, and a negative BVR score means a player is underperforming their contract. Although this doesn't quantify a player's performance like Passer Rating does, it can give us a valuable tool in discussing whether or not they are worth the money they are making.

(transition music)

Alright, let's analyze our BVR scores! First thing's first – we can clearly see that, for the most part, lower paid players have higher value. This makes sense. If a lower paid player can put up the same numbers as a higher paid player, that makes them more valuable. In fact, the top 14 BVR scores are all from players making less than 10 million dollars this year. We've seen some backups come in and play well, like Arizona's Colt McCoy, we've seen some players on their rookie deals play very well like New England's Mac Jones, Arizona's Kyler Murray and Los Angeles Chargers QB Justin Herbert, and we've seen starters on small deals play well like New Orleans' Jameis Winston and Washington's Taylor Heinicke.

The two highest rated players making less than 20 million dollars is Buffalo's Josh Allen and Tampa Bay's Tom Brady. With both players at the top of the league in various statistical categories and only being paid about $10 million each, it's clear to see why they have high value. Other players in this pay grade with positive BVR scores are players like Dallas' Dak Prescott and Los Angeles Rams QB Matt Stafford, who are each considered top 15 quarterbacks in the league.

And then, the negative scores. There are seven quarterbacks making over $25 million this year, and each of them have a negative BVR score. Some of them, like Atlanta's Matt Ryan and Pittsburgh's Ben Roethlisberger – owners of the two lowest BVR scores, have been having abysmal years while being very highly paid. Others, like Green Bay's Aaron Rodgers and Minnesota's Kirk Cousins have been having good years, but the model suggests that that much money just isn't worth it for any quarterback.

And of course, some players who have a BVR score right around 0 and are perfectly paid are Chicago's Justin Fields at 3 million dollars, Detroit's' Jared Goff at $10 million, and Indianapolis' Carson Wentz at $21 million. Wentz has been better than Goff who's been better than Fields, but they are all equally valuable based on where they are being paid.

Let's look at some comparisons across the league to get a feel for how this metric works. Please keep in mind that, at the time of recording and presenting this project, the data accounts for the first 12 weeks of the 2021 NFL season. When you hear this these numbers will be a bit outdated, but the point will still be true, barring some insane future scenario which, in that case, hello and touche. But some comparisons: Matt Ryan and Washington's Taylor Heinicke have had very similar years: Ryan has 2600 yards, 16 touchdowns and 11 interceptions. Heinicke is almost identical at 2600 yards, 16 touchdowns and 10 interceptions. Ryan, however, is being paid $27 million while Heinicke is being paid just 1.5 million. Because of this, Matt Ryan has the lowest BVR score in the league at -7.7 and Heinicke is 7th with a positive 8.9. With identical production and a 25.5-million-dollar reduction in salary, Washington is getting great value out of Heinicke while Atlanta is not with Matt Ryan.

Here's another one: Tom Brady, Baker Mayfield and Jared Goff are each making about 10.5 million dollars this year. Among these three, in terms of passing yards, passing touchdowns and interceptions, Brady is first, Mayfield is second and Goff is third in all three categories. Naturally, Brady has a BVR score of 7.3, Mayfield has a 2.0, and Goff is right at 0.1. Brady and Goff are being paid the same amount, but with 1,000 more passing yards and 20 more touchdowns, it's clear to see why Brady has been much more valuable at the same price.

Now that we have this metric, we in the media can discuss quarterbacks in a different way. The NFL is about building a team within the salary cap, and history has shown that financially balanced rosters are more likely to win it all. Since 2000, when QB salaries really began to start growing, the team with highest paid quarterback in the league has won the Super Bowl just once – Peyton Manning with the Colts in 2009. In fact, that was the only time that the team with the highest paid quarterback even MADE the Super Bowl since 2000.

Quarterbacks are definitely important to the game, and most people will argue that they're the most important, but we haven't seen value in quarterbacks with massive contracts. Kirk Cousins has the highest cap hit in the league this year at $32 million, which is 16% of the Viking's cap space. A player who makes up just 1.9% of the roster is earning 16% of their allotted funds. Teams who give quarterbacks these deals just simply don't have the money to construct a super bowl caliber team around them. This has been a misnomer for years, where people seem to think that a quarterback will simply carry them all the way. In all of Tom Brady's Super Bowls – all 7 of them, he's never been close to the highest paid player that year. The Patriots and the Buccaneers got him for a great value, and then had money to put together great defenses, great offensive lines, and could afford other pieces. The problem the NFL has right now, especially with Patrick Mahomes' half a billion-dollar deal, is that you need more than a quarterback to win, but the quarterback is the one who gets all the money.

It isn't about spending money in the NFL, it's about spending money the right way. The BVR score is now a tool that can be used to determine if a team is spending its money in the right way on it's most important player. One thing to note when talking about these high salaries is also rookie deals. When a player is drafted into the NFL, they are almost always signed to a 4-year rookie deal. If a team drafts a quarterback that gets really good really fast, that team then has a 4-year window where they are getting massive value from that player and they have the funds to buy an excellent team around him, before he gets too expensive. We saw this with the Chiefs going to two Super Bowls with Mahomes still on his rookie deal, and the LA Rams made it in 2018 with Goff on his rookie deal. That Rams team was able to have a massive free agent spending spree and went almost all the way because they had the funds to produce an extremely well-rounded team.

This discussion of value has to take place in our discourse about NFL quarterbacks. Sometimes it doesn't always matter if you have the absolute best player at quarterback, but rather one that you can get great value out of and put a great team around. Joe Flacco won a super bowl with the Ravens in 2013 and he was the 16th highest paid quarterback in the league. The Eagles beat the Patriots in 2017 with a backup quarterback in Nick Foles making 1.6 million and their

starter, Carson Wentz only being the 28th highest paid qb in the league. If you look at the historical trends it's the value that wins championships, not the price tag.

When we talk about free agency, when we talk about trades, when we talk about the NFL draft we have to be sure that we aren't ONLY discussing talent, but rather the team that is getting the most talent for the cost. With the BVR score, we can now reopen the discussion on the superstar philosophy against the egalitarian one, meaning spending boatloads on one guy to carry the team, rather than signing a B to B+ quarterback and giving him a great supporting cast. Hopefully this can spark a new conversation in the NFL. The true formula for success in the league is figuring out the puzzle that is the salary cap, and this metric gives front offices, media members, fans, and even players and agents a better lens when determining quarterback value.

(transition music)

In presenting this new metric, I would be remised if I didn't admit that it is incomplete. It's mathematically accurate, but I still feel like there is some tweaking that could be done. For example, one thing that I left out of the algorithm was fumbles. I felt as though quarterback fumbles are often a team stat, because so often it comes from a strip sack which is often the fault of the offensive line. However, turnovers are turnovers and players like Daniel Jones who fumble a lot should be penalized.

I also wasn't sure what to do with quarterbacks who only start one or less than three times. We've seen some backups this year like Cooper Rush for the Cowboys and Mason Rudolph for the Steelers start once and do decent, but they likely wouldn't be able to maintain a high level of play over a number of starts. Mike White of the Jets came in after an injury to their starter and earned a lot of hype by throwing for 400 yards and 3 touchdowns. His next game he regressed back to the mean in a big way, throwing 4 picks. I considered leaving these guys out of the regression because their good performances on low salaries hurt the value of highly-paid players, but you also pay those players a good amount of money because they're more suited to give you high production for all 17 games in a season. Their ability to maintain production rather than one backup having an okay game should translate to more value, and right now it doesn't

I also considered putting this metric on a 0-100 scale. I actually was able to, by translating the BVR scores into Z-scores and generating percentile amounts. This actually rates each player somewhere from 0-100, and is equivalent to saying "of all theoretical quarterbacks being paid this amount, he ranks in this percentile". Looking at our BVR scores, the highest score is a 16.8 and belongs to Jameis Winston of the Saints on a 2.5-million-dollar deal. Translating that to percentile, his score is a 95, which is much more interpretable for the average fan. Other names we've already mentioned on this ranking system: Taylor Heinicke scored an 81, Mac Jones scored a 77, Tom Brady scored a 76, Baker Mayfield a 59, Jared Goff a 50, Aaron Rodgers a 45 and Matt Ryan a 22. In this rating system a score above a 50 would indicate value and a score below 50 would indicate loss of value. I considered adopting this rating system and I still might. One of my gripes with Passer Rating was its 0-158.3 scale because it's not easily interpretable.

Well, if I said “he has a BVR score of 7” that might not mean much, if anything, so an easily readable scale would likely help the interpretation of my results.

And finally, the next obvious part of this project, is adding a model that answers the question, “what *should* he be paid?” if I am claiming that a player is over or under paid, the next step would then be to predict how much exactly he is over or under paid by and how much he should be worth. To complete this part of the project I’ll flip my x and y axes and run regressions from there, but that will be completed at a later time. I will go back in and edit this podcast once the second part is completed but for now, that remains our next area of inquiry.

(transition music)

That’s it! I hope you enjoyed learning about the BVR Score, also known as the Bagley Value Rating. For a complete look at my project the results, and this this year’s comprehensive BVR scores, I invite you to visit the website I made for this project. The website can be found from my online portfolio, benbagley123.wixsite.com. Or, for a direct URL, put in benbagley123.wixsite.com/bagleyvaluerating. Links to these sites can be found in the bio of this episode.

Before we end this episode, I just want to take one last moment and look at the big picture here. Quarterback salaries are ever increasing, but there is a disconnect between high QB salaries and winning championships. The Bagley Value Rating is an effort to quantify the value of a quarterback based on their stats and salary. This is an important conversation as media members, fans, and front offices because winning football teams are made from value. As this metric continues to be refined and perfected, I hope that you can learn something from it and become a smarter football fan when debating with your friends. In the next episode, we will be talking about quarterbacks out of the NFL draft, and launching a stats project to see if any college stat best indicates NFL success.

Thank you for tuning in to Trade Bait season 3. This was episode 2, the Bagley Value Rating. My name is Ben Bagley, thank you for listening, and I’ll see you next time for episode 3.

Episode 3
College QB Draft Project

Hello and welcome to the third season of Trade Bait. I'm your host, Ben Bagley and thank you for joining me for episode 3. Our first two episodes talked about quarterbacks, and specifically different metrics that exist to analyze their success. Episode 1 talked about passer rating, the official NFL statistic for quantifying quarterback performance. We poked some holes in it and also highlighted other metrics that have done a good job to attempt this quantification, namely ESPN's QBR and Pro Football Focus grades. In Episode 2 we took it a different direction and talked about a quarterback's value rather then his performance. I proposed to you the Bagley Value Rating, a metric that I created to try and capture the value of a given quarterback based on his stats and his salary. Here in episode 3, I want to go down a different route in the conversation about NFL quarterbacks, and I want to go to the place where nearly every NFL quarterback starts their career – the NFL Draft.

(intro music)

The NFL draft is held in late April and early May, and it is a beautiful 7-round battle between all 32 NFL teams to land their next superstar. Teams can trade picks, move up and down in the drafting order, all in an attempt to maximize their haul of amateur talent from college football. As we have mentioned that the quarterback is widely considered the most important position for on-field success, quarterbacks are often the most coveted prize in the draft for teams in need of one. Since the modern draft era began in 1967, 26 of the 54 #1 picks have been quarterbacks. Some qb's, like Troy Aikman, Peyton Manning and Michael Vick go on to change their franchise with hall of fame careers. Others, like Jamarcus Russell, David Carr and Sam Bradford, are infamously known for being busts and quickly fell out of the league.

For so many teams drafting a quarterback in the first round, it is a potentially franchise-altering decision. And, as a great generation of quarterbacks have been ageing and retiring – guys like Philip Rivers and Drew Brees, and soon to be Ben Roethlisberger, Tom Brady and others, teams in recent memory have been drafting first round quarterbacks at historic rates. In 2018, 5 quarterbacks were drafted in the first round, most since 1983. In 2019 3 more were taken, another 4 in 2020, and 5 once again in 2021 – that's 17 first round quarterback selections in the last four years. Some of those players have been exceptional. Lamar Jackson and Patrick Mahomes have each won an MVP, Josh Allen and Kyler Murray look like franchise quarterbacks, and Patriots rookie Mac Jones this year is currently the NFL leader in Passer Rating. Others, however, have been drastically different. Three years after being taken #10 overall Josh Rosen finds himself bouncing around practice squads, Dwayne Haskins was cut halfway through his second season, and Sam Darnold was traded by the Jets and then benched by the Panthers.

So not all quarterbacks that are drafted in the first round pan out to be what the team hoped. Some boom big, some bust badly, and some are just average players. With growing coverage of the NFL Draft, college football, the NFL scouting combine and more, these quarterbacks are being put under the microscope more than ever before – but some teams are still whiffing their

selections. I decided to look at this issue from a statistical perspective, to see if there was something we're missing while scouting quarterback prospects. In this episode, I will present to you my findings from my latest statistical project, attempting to see if any one college statistic is best indicative of NFL success.

(transition music)

The inspiration from this project came from an observation that college quarterbacks come in all different shapes and sizes. There are great running, Heisman-winning quarterbacks like Kyler Murray and Lamar Jackson, there are players like Deshaun Watson and Joe Burrow that dominated college and won the national championship, there's Josh Allen who went to Wyoming and Carson Wentz who went to North Dakota State, and a plethora of guys from college football powerhouses Oklahoma, Alabama and Clemson. But for each of those success stories, there are players from the same backgrounds who couldn't cut it in the league. Johnny Manziel was an electrifying athlete who won the Heisman and found himself playing in the Canadian football league a year after being drafted. Tim Tebow, Matt Leinart and Vince Young were all quarterbacks in the 2000's whose teams won dominant national championships, but each busted in the pros. There are other guys who went to smaller football schools too like Blake Bortles at University of Central Florida and Jared Goff from Cal, and they haven't found their spot in the league along with a large number of players form some of those powerhouse programs.

So what's the difference? What's the secret sauce that makes a quarterback good in the NFL? With the successes and failures we've seen from quarterbacks from all range of backgrounds, how can we determine which new college prospect will find NFL success? That was my driving question in this project. I wanted to see if there was any one college statistic that most predicted a high NFL passer rating and if so, which one and how strong of a predictor was it? If there is a college statistic that successful NFL quarterbacks share, then that is an area that front offices and media members should look at and adjust our evaluations accordingly. If there isn't, then we'll have to revisit the evaluation process as a whole to predict success.

The media, both big and small, has a very strong presence in the NFL draft process. NFL front offices might not admit it, but the work of so many personal private scouts and the growing presence of advanced analytics has been drastically influencing the NFL Draft landscape for the last decade. Teams are imitating others, looking for similar success, rolling the dice quite often, and looking for a glimmer of light through the hit-or-miss draft process.

In this project, I took every college quarterback drafted in the first round since 2005, identified 25 different college variables that I thought of as important, and ran a lasso regression to identify the strongest predictors. The results might surprise you.

(transition music)

In setting up this project I took every quarterback drafted in the first round since 2005 – all 43 of them -- and identified 25 college statistics: 21 on-field stats and four off field. I wanted to include a wide range of variables to account for as many things we could quantify as possible. Some quarterbacks are 4-year starters, some only have one year of statistics to their credit, some ran a lot in college, etc. For my variables, I included some career passing statistics like total yards, total attempts, total completions, yards per attempt, career touchdowns, career interceptions and college passer rating. I also included rush-friendly variables like total rush attempts, total rush yards, and rushing touchdowns. I included total wins, total losses, and career winning percentage as well. Then I also used four off-field statistics: the number of undergraduate students at their university the year they graduated, to see if the size of the school predicts anything, the amount of years they were on a roster and eligible to play (which excludes years where they transferred and ineligible to play for that year), their Wonderlic score, which is basically an IQ test for quarterbacks, and what number pick they were in the draft. The final variable is the response variable, NFL Passer Rating. In running this regression, I tested each of these variables together to try and see which one, if any, was the strongest predictor of NFL passer rating.

The statistical process that I ran to do this is called a lasso regression. This is a type of linear regression that uses shrinkage and variable selection. Shrinkage is where data values are shrunk towards a central point, and if theyre shrunk closely enough to zero, they are essentially eliminated from the model. By seeing which variables are significant enough predictors to be eliminated from the model, the lasso essentially suggests which variables are strong predictors, and it explains how much of a predictor it is. This process is an estimation, so every time you run it, depending on the correlation your variables, the final model suggestion could include different variables and suggest different significant predictors. Also, in this process, the data is also standardized so everything is on the same scale. A player could have 20 rushing touchdowns and 10,000 passing yards, so we put those on the same workable scale.

Our data is highly uncorrelated, and this makes sense. Knowing how many total passing attempts a player had in college will likely tell you very little about their Wonderlic score, and knowing their college win percentage will likely tell you very little about the amount of rushing touchdowns a player had. Because of this, our lasso suggested different models with different strong predictors every time I ran it. So, I ran it 100 times and averaged the results. The goal of this process was to see which college statistic was included as a significant predictor variable of NFL Passer Rating most often in these lasso regressions.

Our results, are shaky. The lasso spits out two results: how often a predictor variable was chosen as significant, and how much an increase or decrease in that predictor changes the response variable. Again, our predictors are college stats and our response is NFL passer rating. The predictor variable chosen at the highest rate was rushing touchdowns, chosen 29% of the time. The next highest was college interceptions, chosen 19% of the time, Wonderlic score, chosen 19% of the time, rush yards per attempt, chosen 15% of the time and school size, chosen 14% of the time. Interesting to note that only one passing statistic was in the top 5 most selected predictors, possibly suggesting passing stats aren’t as indicative of NFL success as some

other stats. The least chosen variables were winning percentage, just 2% of the time, completion percentage, just 3% of the time, passing touchdowns, just 3% of the time and total passing attempts at just 4%. Interesting that we see some of these very standard football statistics not being seen as strong predictors, just showing how large the variance in these stats are in first round quarterbacks.

The next thing that the lasso spits out are beta coefficients. These coefficients tell us, for every one unit increase in a predictor variable, the response variable changes by that much. Here is where we can start to see the results of this stats project: that there really isn't a college statistic that tells us much about NFL success.

The highest beta coefficient corresponds to our most chosen predictor, rushing touchdowns. The beta coefficient is 0.0842. That means, for every additional rushing touchdown a quarterback has in college, their expected NFL passer rating increases by 0.0842. That is a very small number. In fact, that means that a quarterback would need 12 more rushing touchdowns in college than another quarterback to have an NFL Passer Rating that is just one point higher. Our next most common predictor was interceptions, and its beta coefficient is 0.0505. So, for every additional interception a quarterback throws in college, their expected NFL passer rating increases by 0.0505. Intuitively it seems backwards that more college interceptions would increase your expected NFL passer rating, but that's what the data tells us. So in this category, a qb would need to throw 20 more interceptions than another qb to have a NFL Passer Rating that is one point higher. I could keep going with this but you can see that these beta coefficients are extremely small, meaning that, for the most part, these are all very weak predictors of NFL passer rating.

Let's contextualize this, and you'll quickly see why there isn't a specific go-to stat that best works as a predictor. Let's look at rushing touchdowns, which the lasso said was our strongest predictor: the quarterback who had the most rushing touchdowns is Tim Tebow, with 57. His NFL passer rating is actually one of the worst in the data set, at 75.3. Remember, passer rating is on a scale of 0-158.3. The quarterback with the next most rushing touchdowns is Lamar Jackson with 50, and he has an excellent NFL passer rating at 101.7. Third, however, is Vince Young and he has a 74.4 NFL Passer rating. The next four players in this category – Robert griffin the third, Johnny Manziel, and Jake Locker – all have an NFL Passer Rating under 85. So if you're in the draft and you choose the quarterback with the most total rushing touchdowns, you could either get Lamar Jackson or Tim Tebow, which is a drastic difference.

Let's go to a different variable though, and our next most common predictor was college interceptions. Players in the top 10 in this category who have an NFL passer rating over 90, which is good, are Matt Ryan, Deshaun Watson and Matt Stafford. However, the quarterback who threw the most interceptions is Brady Quinn, who owns the second-lowest passer rating in this data set at 64.4. Other players in the top 10 are subpar NFL quarterbacks Josh Freeman and Christian Ponder. So, if youre in the draft and you choose the quarterback with the most interceptions, you could either get Matt Stafford or Brady Quinn – another drastic difference.

The problem is, is that in any college statistic you choose, the top 5 quarterbacks in that stat will include an NFL great and an NFL bust. Most Passing Yards? Brady Quinn is 3rd and Patrick Mahomes – who owns the highest career NFL passer rating of all time – is behind him at 4th. Passing touchdowns? Baker Mayfield and his 89.9 NFL passer rating is first, but Matt Leinart and his 70.2 NFL passer rating is 3rd. Highest winning percentage? Cam Newton is first but Dwayne Haskins is 3rd. School size? The quarterbacks from the biggest schools were Blake Bortles, Dwayne Haskins and Johnny Manziel, each with a passer rating under 80. The quarterbacks from the smallest schools were Jay Cutler, Daniel Jones, Andrew Luck and Brady Quinn, which is another collection of quarterbacks with high variance in their NFL performance.

The short answer is this, and if you got lost in the numbers here's our conclusion: for any of these college statistics, there have been first round draft picks that have been very good in that category that have both boomed and busted in the NFL. No matter where you look, college statistics largely don't tell you much about that player's success in the NFL. There have been so many great statistical college players that have boomed in the NFL, so many great college players that have busted, and vice versa. There have also been so many statistically average college quarterbacks that have done really well in the NFL, and so many that have done poorly or just average.

Here's what this regression told me: No, quite simply, there is not an NFL statistic that can accurately predict NFL success. The career trajectory of a top draft pick cannot be accurately determined by his college statistics, but rather a number of other non-quantifiable variables. There's a large mental aspect to the game, where some players are just built differently than other. Some rookies get drafted to great situations with a good head coach and solid pieces around them which lead to success, and other rookies get drafted to horrible situations where they couldn't overcome the dysfunction of the franchise. And vice versa, too: there are rookies drafted into great situations who don't pan out well, and there are rookies who have brought a garbage fire of a team back to greatness. But unfortunately for the statistician in me, this success or lack of success is not predicted by their college stats. The stats can certainly help predict success, but this project shows that there is **much** more too it at the NFL level. College numbers simply do not accurately translate to NFL success.

(transition music)

So, what does this mean for us in the NFL media? Well, it should largely influence our coverage of the NFL draft. NFL fans watch coverage of the draft and they pay attention to the scouting process because they want to know if their team's quarterback selection is going to be great or not. So, if we in NFL media are only using college stats to analyze these players, we are doing fans a large disservice by providing them with hollow information. There is more to NFL success than college stats, so we need to acknowledge that in the draft process.

This means we need to break down film more. We need to interview and dissect the mental aspect of players more. We need to talk about the current state of the NFL team that is

drafting, and if they are in a situation that would be cohesive for a rookie to find success in. Just choosing the best statistical player and plugging them into any roster won't predict success, but it's rather about the best specific player for the best specific team at the best specific time. Every player is a completely unique case. What this stats project really tells is that finding a franchise quarterback in the NFL draft is incredibly hard. Scouting and predicting success at the next level based on statistics is all but a total shot in the dark. Either there needs to be more advanced metrics in college football, or we need to analyze players in a different way.

We need to be aware of this phenomenon as active members of the NFL media to make sure we aren't doing fans a disservice. We want to be as accurate as possible in our predictions, so it is important to know what part of the prediction process is actually inaccurate. The smarter we can be as NFL media members, the more engaged the fans will be, the smarter the fans will be, and the better our NFL discourse as a whole will be. I love the NFL draft, I love scouting and I love trying to predict the next GOAT, but it is time to formally acknowledge the fact that drafting a quarterback in the NFL draft is a complete and utter crapshoot.

(transition music)

That's it for the third episode of the 3rd season of Trade Bait. In the next episode, we'll shift our conversation from quarterbacks to the rest of the offense. Running backs and wide receivers, in specific, are just as worthy of this kind of analysis as quarterbacks are. In episode 4 we will break down the main stats for these skill positions, and examine their pros, cons, and offer suggestions for improvement. Now that you've heard this episode, I hope your next NFL Draft season brings you a reason to look a bit more critically about the top quarterback prospects. Every team loves landing their superstar, but the process is more difficult than even the best scouts could predict. Thank you once again for listening to this episode and the entire third season of trade bait, my name is Ben Bagley, and I'll see you next time.

Episode 4
RB, WR, TE Statistics

Hello and welcome to Trade Bait season 3, episode 4. I am your host, Ben Bagley. In the first three episodes we focused exclusively on the quarterback. We broke down Passer Rating, learned about the Bagley Value Rating, and discussed difficulties in the scouting process of college quarterbacks, and how their college stats are not necessarily strong predictors of NFL success. In this episode, I want to give some love to the rest of the offense. We're going to discuss the current statistical state of running backs, wide receivers and tight ends – often referred to as skill positions. Although I don't have a new metric to propose or a groundbreaking stats project to present, there are plenty of statistics to unpack and nitpick. Let's get it started with the pass catchers, we're talking about wide receivers.

(transition music)

Statistically speaking, wide receivers are underserviced. We've been talking about all kinds of advanced metrics that exist for quarterbacks, but there isn't a lot of advanced stats in the main stream for wide receivers. Their main statistics, as they've been for so long, are catches, yards, and touchdowns. In the last two decades we've began using stats like yard after catch more, and NFL next gen stats have introduced us to deeper statistics like speed with the ball, average yards of separation from their defender, and targeted air yards. But still, there is a lot that is left to be desired in the effort to quantify skill.

The first thing that needs to be mentioned is the randomness of their statistics. On any passing play, the quarterback will have one to five have possible pass catchers on the field. Sometimes teams design an obvious play to one receiver like a screen, but most often there are multiple players running routes and the quarterback is responsible for choosing his target. This is the first problem with wide receiver stats. What a wide receiver has control of on the play is essentially only themselves. They have control over how well they run their route, and that's it. They have no control over the defense, and who is covering them or if it is man-to-man or zone coverage. Man to man means a receiver needs to beat their defender to get open, and zone means they need to weave between defenders and find an open area. They also have no control over if they even get the ball or not. No matter how open or how covered they may be, it is the quarterback's decision of when and where to throw the ball. A wide receiver could be wide open and the pass comes to him, or he could be wide open and the ball isn't thrown his way. He could also be covered and the ball is still thrown to him, or he could be covered and the ball isn't thrown his way.

The thing that's been bothering me down this line of thinking is that often we use catches, targets and catch percentage to quantify wide receiver success. Catches is obviously how many times they caught the ball, targets is how many times the ball was thrown their way, and catch percentage is the percentage of targets they caught. But the problem is that targets is not a wide receiver stat, it's a quarterback stat. And down that line of thinking, the amount of catches a receiver has is almost entirely dependent on how many times the quarterback threw it to

him. And if he's covered very tightly all game, and he gets some targets but no catches, his catch percentage will go down even if the ball shouldn't have been thrown his way.

So when talking about wide receiver stats, this is the first thing we need to be aware of: their numbers are extremely dependent on the quarterback and these basic stats do not do a good enough job of quantifying their abilities. Just because one receiver has more catches as another does not necessarily make him better, but most likely means that the quarterback threw it to him more often. And similarly, more targets does not make a receiver better, but rather means the QB just threw it to him more. So using those stats as the main way of determining if a wide receiver is good or bad is antiquated.

In order to take this a step further, lots of statistical data is needed. First, since we want to be able to quantify success beyond these basic stats we have now, we will need to access player tracking data to see what the receiver is doing regardless if they have the ball or not. A stat that I think would much better evaluate the abilities of receivers would be average separation and open percentage.

The first one, average separation, has already begun to be tackled by NFL next gen stats. This stat is the distance (in yards) measured between a receiver and the nearest defender at the time of catch or incompletion. In theory, this tells us how good of a job the receiver did at getting open. A higher average separation distance means that the defenders are farther away from the receiver when the ball gets there, which should mean the receiver has a better chance to catch the ball. The problem here is that this metric doesn't account for the defensive coverage. If a team is in zone coverage – and there are dozens of different zone coverage schemes – defenders will each cover a specific area of the field. If a receiver enters their area then its their responsibility to cover them, and if the receiver leaves that zone then it's the next defender's responsibility. If it is man to man coverage, then each receiver will have a specific defender assigned to cover them and they will shadow the receiver's movement on the field. With zone coverage, there is not a defender glued to a receiver like there is in man coverage. The philosophy behind zone is to have more defenders around the receiver when the ball is caught to rally and make the tackle, or to confuse the quarterback by just causing mayhem in coverage. But for zone, the defender tries to be as close as possible to the receiver to not let them catch it at all.

So, naturally, receivers can find high separation from their defender much easily in zone coverage. Because nobody is assigned to be glued to him, nobody is and it's much easier to get open. Man coverage typically makes it much harder for a receiver to get open, as they need to use fancy footwork or blazing speed to separate themselves from their guy. So for this statistic of average separation, we need to differentiate which players are getting high degrees of separation in zone coverage vs man coverage, because that tells us something deeper about each receiver. Are they consistently running great routes to shake man coverage, or are they consistently finding soft spots in the zone? Getting separation regardless is a positive, but it's important to know how they are getting their stats.

From here, I want to know a receiver's open percentage. Completely independent of the quarterback choosing to throw it or not, I want to know how often the receiver is open and available to his QB. To do this, we'd need to use NFL next gen data to measure the average degree of separation divided by the number of routes ran. To take it further, we could break up routes into short, intermediate and deep zones of the field to determine their openness percentage on different kinds of routes, and we could also separate it in man vs zone coverage. This statistic, knowing how often a receiver is open, would be much more useful, in my opinion, than knowing how many targets they had. Targets is a quarterback stat, openness is a wide receiver stat. The goal here is to assign the receiver a stat that they can actually control. I don't really care who had the most targets, I care who is the best at getting open. From here we could do a lot: how often the ball is thrown to them when they're open, how often the quarterback is finding open receivers, et cetera.

In my research on this topic, I actually stumbled across an article FiveThirtyEight.com from August of 2020. FiveThirtyEight is a journalistic stats website that focuses on politics and sports, and they are always trying to create new metrics to quantify all this information in creative ways – if you're from FiveThirtyEight.com and listening to this podcast, an internship here would be right up my alley, wink wink nudge nudge. But anyway, FiveThirtyEight created a statistic to do just this. Using NFL Next Gen Data, available statistics and considering a lot of situational football as well, they created what they called "Separation over expected", or SOE. They used data to create a model which predicts the average expected separation on any given route in any given situation by any given average receiver. Then, using a real receiver's separation data in that specific situation, they determine how much more or less separation they had then what was expected. I read through this article and they actually did an exceptional job with the metric, I will include the link to it in the bio of this episode. They also included depth of pass so we were able to see who was the best at getting open on short, intermediate and deep routes. Their project answered a lot of these questions I proposed, which was very fun to read through.

Now, as members of the NFL media, how can we use this? I recommend getting away from the use of catches and targets in our evaluation. Of course, it's always going to be relevant to know who leads the league in catches and targets and yards and touchdowns and everything, but we need to include this kind of metric so we can evaluate a receiver on the things that he can actually control. Receivers who put up bad with really poor quarterbacks don't deserve as much blame as their statistics suggest, because those number are often so dependent on the quarterback's ability to get them the ball. One player that comes to mind in this conversation is Bears wide receiver Darnell Mooney. I remember after his rookie year, where he put up average numbers, there was an entire highlight reel of him running beautiful crisp routes and completely shaking his defender and then the ball was overthrown from Mitch Trubisky every single time. Mooney now isn't considered a great receiver, but the tape, which doesn't translate to stats because of his quarterback, says he might just be one of the best in the league. As members of the NFL media I just want us to be able to acknowledge the skill of players who are quietly torching defenders even if their stats don't say so, and I want us to stop crowning the best players as the ones who get the most targets. As a media outlet,

FiveThirtyEight has done an excellent job of making and explaining metrics like this, but it needs to work its way into the mainstream so we can better evaluate the wide receiver position.

(transition music)

Next, I want to talk about tight ends and the unique position they play in the offense. Tight ends can be a pass catcher like a wide receiver, or they can be a blocker and function as part of the offensive line. For the receiver part of their position, everything we just talked about holds true: targets and yards vs openness percentage and separation over expected would be a valuable improvement in their statistical dictionary. But because of the duality of their position, we need another stat for tight ends.

With the exception of outstanding talent like Travis Kelce and Rob Gronkowski, tight ends typically have much smaller statistics than wide receivers because they spend more time blocking. Tight ends also tend to come in three different varieties: vertical threats who are essentially just big wide receivers who are more skilled at catching than blocking, blockers who are more skilled at blocking than catching, and hybrids who can do both well. What I want, though, is a metric that can marry their blocking stats with their receiving stats. In the NFL right now, two of the top tight ends are Travis Kelce of the Chiefs and George Kittle of the 49ers. Kelce has some of the all-time best receiving stats for a tight end, but Kittle is an exceptional blocker while still having great receiving stats. I want to compare the entirety of their skillsets, and because of their multi-faceted role in the game, I think it's important to look at more than just receiving stats to determine who is better all-around at the position. But if I want to add receiving statistics to blocking statistics to create a comprehensive metric for total tight end production, we first need to talk about blocking stats.

The problem that exists right now is that blocking and offensive line statistics as a whole hardly exist. Really only with the introduction of NFL Next Gen stats in 2017 and through work of private stats companies like Pro Football Focus and Football Outsiders have we only recently begun to create offensive line statistics. We'll talk about this more next along with our discussion with running backs, so for now I'm going to keep this brief: but we need a way to evaluate how much a specific blocker adds to the run or pass blocking on the play. But that is an incredibly difficult thing to quantify. As with the problem with wide receiver stats, in the NFL right now we only really keep stats of players who possess the ball, everything else is mostly just white noise. And this makes sense, because imagine how hard it is to put a numerical statistic for every blocking player on every play. With five offensive linemen and a tight end moving on a 2-dimensional plane against 11 potential defenders they need to block, these potential statistics almost turn into a multivariable calculus problem. If I want to know how many yards a specific blocker added to a run, I need to know how many yards the running back was expected to gain and how much one blocker personally added or subtracted to what he actually gained by how far he moved the defender, what direction he blocked him in, the time that he was engaged with the blocker, the hole the runner chose to hit, et cetera. There are so many non-numerical variables included that it very quickly gets away from my style of statistics.

We need motion tracking data and advanced models to create a metric like rush yards over expected, which is a real stat, but we'll talk about that with running backs.

So for now, let's table this discussion on tight ends but keep in mind the point I made earlier: for a more complete statistical evaluation of a tight end, we need to both quantify their blocking stats and add them to their receiving stats to create a new metric that fully quantifies everything that the unique position of tight end does on the field.

(transition music)

Finally, let's talk about running backs. Running backs have been an interesting topic of conversation in recent years in the NFL because they've just seemed so replaceable, no matter how good they seem to be. Remember when Rams running back Todd Gurley had two outstanding seasons in 2017 and 2018, being named all pro and leading the league in total touchdowns in that span? Two years later, at age 27 he can't find a job in the NFL. It's not his fault, this has just been a growing narrative in the NFL: running backs are replaceable, so don't give them big money.

I started noticing this a few years ago, but my eyes were really opened this year while I was playing fantasy football. There are so many excellent running backs in the league that put up huge statistics, and many of them got hurt for stretches this year. But, curiously enough, their backup was able to fill in flawlessly. Viking's star running back Dalvin Cook missed stretches of games twice this season; his backup Alexander Mattison was excellent as a replacement gaining over 90 rushing yards in all three of his starts with two touchdowns. The Browns have two star running backs, Nick Chubb and Kareem Hunt. They both missed a game in Week 7 and the fourth string backup D'Ernest Johnson came in and promptly put up 146 yards and a touchdown. When Chubb and Hunt were injured again in Week 10, Johnson came back in to rush for 99 yards and he had 7 catches for 60 yards. When NFL rushing king Derrick Henry was lost for the season due to an injury, the Titans put in D'Onta Foreman and Dontrel Hilliard, and both former practice squad players had over 100 yards rushing in the SAME GAME against new England. Other backups just this year that had this kind of success also include a pair of rookies in Patriots' Rhamondre Stevenson and Panthers' Chuba Hubbard.

So if this in an increasing trend across the league, a lot of questions pop up. The first, to me, is the question of the offensive line. The offensive line is almost wholly responsible for the running game. They make the blocks, open up holes for the runner; they set up the play, anything else the runner does is extra. With a good offensive line, you're likely to see good rushing numbers. So here lies my question: If a backup is able to come in and put up just as good of numbers as the starter, how much of the starter's perceived greatness was actually just due to the offensive line? If seemingly anybody can be plugged right in and put up big numbers, is it worth even having an elite but expensive running back on your roster? What we need to know here is, once again, how much the offensive line is contributing to the rushing numbers vs how much a specific running back is.

Judging the offensive line as a whole is much easier than just one specific player like a tight end, and people have begun to attempt to perfect these kinds of stats. My favorite stats on the market right now come from Football Outsiders.com. They have all 32 offensive lines ranked based on a number of film-based statistics, starting with adjusted yards per carry. They take each carry a running back has and how many yards he gained, and then they assign responsibility to the offensive line based on this scale: loss on the play, 120%, 0-4 yards is 100%, 5-10 yards is 50% and 11+ yards is 0%. They also have second level yards, which is yards the running back earned 5-10 yards past the line of scrimmage, and open field yards which is yards the running back gained 10+ yards beyond the line of scrimmage. The goal with these numbers is to see how much of the run the running back was responsible for vs the offensive line.

For example, A team with a high ranking in Adjusted Line Yards but a low ranking in Open Field Yards is heavily dependent on its offensive line to make the running game work. A team with a low ranking in Adjusted Line Yards but a high ranking in Open Field Yards is heavily dependent on its running back breaking long runs to make the running game work. However, it is important to understand that these ratings only somewhat separate the offensive line from the running backs. A team with a very good running back will appear higher no matter how bad their line, and a team with a great line with appear lower if the running back is terrible.

So, this an effort to quantify the statistical value of a running back through the lens of the offensive line. What if we want to do it from the eyes of the running back? Two statistics come to mind, and one of them already exists: rush yards over expected. This metric has been created and used by professionals at NFL next gen stats and pro football focus, and it is a doozy to calculate. They get in the nitty gritty and calculate How many rushing yards a ball-carrier is expected to gain on a given carry based on the relative location, speed and direction of blockers and defenders. They then generate a model that produces expected rushing yards for any given play in any situation. From there, they can plug in how many yards a runner actually gained and claim that it is more or less than the expected number of rushing yards. This is a way to explain how much of the running game's success or lack of success can be attributed to the running back, but it doesn't explain this phenomenon of backups performing at such a high level.

For that, I'd like to propose the idea a new metric. In baseball they have Wins Above Replacement, WAR, and I'd like to propose a similar metric for running backs, RBWAR. In baseball, WAR measures a player's value in all facets of the game by deciphering how many more wins he's worth than a replacement-level player at his same position. I'll save you the calculation because it's a doozy, but the metric quantifies each players value in terms of wins. And they do it through the eyes of a replacement-level player and I love that term. In my examples earlier, those were replacement-level players coming in and putting up great numbers. So, if a starting running back's numbers aren't that much better than a replacement, his value should go down. The other thing I love about this metric is that it quantifies success via wins. The stats don't matter as much, but rather it's about who can we put back there that will generate the most wins for the team. I think we can also add salary and value into this metric like we did for the Bagley Value Rating as well.

I don't have the means to create this statistic right now, but I think it can be an important piece of the running back puzzle. We want to maximize efficiency and production all based on salary, and much more is needed when talking about the run game. Just as receivers are dependent on the quarterback to get their numbers, running backs are dependent on the offensive line to get theirs. And if you have a great offensive line, it will be much easier for a replacement-level running back to come in and have success. If you don't, you might not find running success at all. I just think we need to incorporate a more holistic view of the run game and figure out a way to assign proper credit where it is due, either to the running back or to the offensive line.

(transition music)

Okay, so let's take a minute here and talk about this through the eyes of the media. Why does this kind of conversation matter? Well, the common theme through this discussion is assigning proper credit where it is due. Football stats are so dependent on having the ball in your hand and, since the quarterback has the ball in his hand every play, he gets the most attention and is the main person we assign credit to. But to offer a quarterback-independent view of other players, we need to be more diligent in our statistics.

Only a very small part of what happens on the field is actually reflected in the current statistics that we use. With 22 players on the field moving at the same time in a beautiful and brutish ballet, there is a lot more to the game than we talk about in the media. I want us to be able to become better evaluators of every position, and I really want us to be able to provide more context when we say "this player is better than that player". If we say this wide receiver is better than that one, I want to see their openness percentage and average separation, not their targets and catches. If we say this tight end is better overall than that one, I want to see their contributions in both the pass and run game combined into one metric. If we say one running back is better than the other, I want to see his numbers independent of an offensive line and in terms of wins above replacement. I want our stats as a whole to just go deeper.

Why? Honestly, to better the NFL media landscape as a whole. To give reason and context behind opinions and claims we make as fans and media members. This episode and these suggestions are more or less a plea to become smarter in our evaluation of NFL talent. There are brilliant people working in all aspects of the game, but our discussion of players is still so basic and antiquated. With these kinds of proposed metrics, and with the increasing possibilies of advanced metrics with next gen data and more, we have the power to improve our collective NFL intelligence in a big way. And with that, these numbers can even start to be used on shows that deal with legalized sports betting or fantasy football. If we can better evaluate a single player at a position based solely on his production, we can start to better predict which players will boom or bust any given week and potentially give fans an edge in this category. Going deeper into these statistics and drawing out this advanced data can only help our current state of discourse. Isolating quarterback vs wide receiver vs tight end vs running back vs offensive line allows us to see what is truly going on on the field, and lets us talk about the sport with more accuracy, intelligence, and credibility.

(transition music)

That's it for this episode of Trade Bait, season 3. Although I didn't present any new metrics, I hope this conversation opened your eyes to the next level of stats that we can take in the NFL. As a fan and as a consumer of the NFL, watch the game more critically. Begin to question it in the way that I do. The main way that these kinds of stats will gain relevance is if there is a strong demand for them. So my parting request for you from this whole season is to request more. Want more. Demand more. Don't just accept stats at a surface level, figure out what they are telling you, what's missing from the story, and what you need to complete the narrative. This has been episode 4. For our next and final episode, we will be interviewing a member of NFL network for a fascinating conversation about the current state of stats in the media. My name is Ben Bagley, this is Trade Bait, thank you for listening and we'll see you next time.