CLUTCH FACTOR: A NEW WAY TO MEASURE HOW NFL QUARTERBACKS PERFORM UNDER PRESSURE

CHARLIE HUME AND MAX FUNG

Abstract. In our project we seek to investigate which quarterbacks in the National Football League today are the most "clutch." Broken down into two different sections, we first develop a new method for grading quarterback performance, then apply that system to specific late-game scenarios to create a final "clutch" ranking. Our ranking, called Quarterback Performance Index (QPI), revitalizes the antiquated passer rating system which has been in use by the National Football League (NFL) for over 50 years. This system accounts for statistics important for a quarterback in today's game, such as passing yards per attempt and passing touchdowns per attempt. By using some of the greatest and worst quarterback performances of all time, we developed a matrix system assigning these performances quarterback grades. Then, by utilizing a Moore-Penrose Pseudo-Inverse we use linear regression to find coefficients for each statistic in our rating. Then, using this rating system, we compared a quarterback's QPI during important late-game scenarios to his QPI in other situations to find his Clutch Factor (CF). Finally, by ranking each quarterback based on his CF, we were able to identify Desmond Ridder and Kenny Pickett as the most "clutch" quarterbacks in the 2023 season.

1. Introduction

In football, there is no position more valuable to the team than a quarterback. As the player with the ball most frequently in his hands, the quarterback has the most individual impact on whether a team will win or lose a game. In particular, in "clutch" situations where games are close and time is winding down, quarterback performance becomes vital. As the NFL has grown in popularity, there has been a natural interest in how exactly to rate how well a quarterback performs in any given game or over the course of a season. However, many of these ratings are either outdated or value the wrong statistics to correctly evaluate a quarterback.

Clutch is a measurement of whether a player steps up in the game's biggest moments or cracks under the pressure. The problem with all the current models that set out to rank the most clutch NFL quarterbacks is that they simply compare the stats of different quarterbacks in scenarios they deem to be most pressure-inducing. This is an issue because the ideas of clutch and sheer talent are obscured. How do we know that a quarterback is playing well late in the game because he is clutch versus him just being good? This is why we tend to see the league's top quarterbacks dominate the current clutch rankings as well. To find out which quarterbacks are truly the most clutch, we need to remove overall talent and skill from the equation. We want to know if a quarterback's typical gameplay is elevated or lowered during crunchtime.

Our goal is to create a new metric called Quarterback Performance Index (QPI) that fits with the modern state of football to properly rate today's quarterbacks. Then, using

Date: December 27, 2024.

that metric, we will compare quarterback performance during clutch situations to overall performance. Finally, we will develop a ranking of which quarterbacks truly are the most "clutch." The results of our work are not what most football fans would expect.

First, we need to create a Quarterback Performance Index (QPI) which will appropriately grade a quarterback on their performance in a given quarter of a football game. To accomplish this, we must select the most important statistics to use as variables that encapsulate a quarterback's performance and implement them into our QPI formula. These statistics need to fairly grade all quarterbacks regardless of their play style; in particular, they must guarantee that rushing quarterbacks are given enough credit for their production on the ground, as well as, through the air. These statistics must be given proper weights in order to output the most accurate rating given any stat line.

After our QPI formula is complete, we then must create a formula to calculate Clutch Factor (CF). CF will compare average QPIs in the first three quarters of a game to QPI during the 4th quarter of a game within 7 points. This will give us a tool to compare quarterback performance during "crunchtime" to average performance. Finally, we will create a ranking of 2023's top 32 quarterbacks' CFs to determine who truly is the most clutch quarterback in the NFL.

In Section 2 we overview related work, and in Section 3 we describe our method. We share our results in Section 4, evaluate the strengths and weaknesses of our method in Section 5, and conclude in Section 6.

2. Related work

2.1. **NBA Clutch Ranking:** There are similar projects to ours done in other sports as well. In [5], Schuhmann tries to identify who the most "clutch" players were halfway through the 2023-2024 NBA season. This article looked at several different factors which could be used to describe clutchness to create their ranking. First, they measured total "clutch points" scored, followed by "clutch accuracy." Both of these statistics measure performance in the last five minutes of games within five points. Other statistics measured include total baskets made within the final minute of close games, as well as assists and blocks. Each category has a ranking attached to it, with each one containing a variety of different players. We use a similar approach of looking at a variety of different statistics, but instead we use them all together to create one compiled rating system.

2.2. **Comparing Clutch Performance in Tennis:** Jetter,[3], looked into whether tennis players play better or worse in the biggest moments. He determined the most important tournaments to be the four Grand Slam tournaments: the Australian Open, French Open, Wimbledon, and the US Open. Jetter also considered tie breakers and decisive sets to be more pressure inducing, similar to our decision to make the fourth quarter within 7 points the criteria for our data. He used logic regression to estimate the likelihood of a tennis player to win their match. In order to calculate our QPI, we use multiple linear regression.

2.3. **Passer Rating:** As we look for the best way to measure quarterback performance, we first examine which stats and metrics are readily available for us to use and learn from. In [4], a blogger going by rufio explains some of the major pitfalls of a widely used measurement of quarterback performance called passer rating. One major flaw is that the four components of passer rating all have limitations on how high of a score a player can receive. This

is problematic because it doesn't allow for record breaking performances to receive higher passer ratings. Another problem with passer rating is that two of its four components are yards per attempt and completion percentage. This is problematic because it's essentially factoring in completions twice. Furthermore, he explains how completion percentage is overvalued to the point where completing a pass for zero yards is the equivalent to throwing for 20 yards. He also claims that interceptions are worth about 100 yards lost, when they should only be worth about 45 to 60 yards lost. Lastly, he points out how touchdowns are also inflated in passer rating because they are worth approximately 80 yards, which is far too high. We will adjust the weights of passing yards, interceptions, and touchdowns to more closely fit these more accurate estimations.

2.4. **NFL Passer Score:** We also explored better alternatives to passer rating, such as Fromal's passer score[2]. Like rufio, he also identifies many of the same problems with the passer rating statistic; however, Fromal took the initiative to correct its errors. He started by addressing the bounds of each component, allowing for record breaking performances to receive a score of 100 or higher. He then created the formulas for each component, comparing the quarterback's performance to the all time average for that particular stat. Players who outperform the average receive a positive score for that component, while players who underperform compared to the average receive a negative score. This method for illustrating performance is very easy to understand; however, it still uses the same four components as passer rating, which we already discussed to be flawed. We believe applying this logic to a new set of components that consider a different collection of stats would allow us to create a highly relevant and intuitive way to measure quarterback performance.

2.5. **CBS NFL Clutch Ranking:** Clawson, [1], ranked the top 10 clutchest quarterbacks at the beginning of the 2024 NFL season, based on two key stats: conversion rate during important late-game plays such as game-tying drives or overtime, and expected points added (EPA) per play in these situations. EPA is found by looking at the outcome of every play, and calculating how much of an impact that play had on the points scored on that drive. Clawson chose to equally value all games regardless of stake, including playoff games. Our project will differ significantly from [1], even though the intended outcome is roughly the same. We will use a completely different method of ranking quarterbacks, comparing their performance in crunchtime to their play in other periods instead of against other quarterbacks. In fact, many of the quarterbacks ranked among the most clutch on this list are not highly ranked on ours and vice versa.

3. Our Method

Our project has primarily two major parts: Creating our quarterback rating system (QPI), and implementing that rating to rank the "clutchest" quarterbacks. QPI is primarily meant to supplant the "passer rating" system, which was adopted over 50 years ago. We used a similar system of taking different metrics and scaling them, but modernized the statistics taken and overall method used. There have been several developments in the game, such as the rise in prominence of rushing quarterbacks, which warrant consideration when evaluating a quarterback's performance. The six statistics we used are, in order, rushing yards per attempt, passing yards per attempt, touchdowns per attempt, interceptions per attempt,

fumbles per game, and rushing touchdowns per game. Our final equation will take the general form:

$$\begin{aligned} QPI = x_1(\frac{\text{rush yards}}{\text{attempt}}) + x_2(\frac{\text{pass yards}}{\text{attempt}}) + x_3(\frac{\text{passing touchdowns}}{\text{attempt}}) \\ -x_4(\frac{\text{interceptions}}{\text{attempt}}) - x_5(\frac{\text{fumbles}}{\text{game}}) + x_6(\frac{\text{rushing touchdowns}}{\text{game}}) \end{aligned} \tag{1}$$

We created a system of equations that contain real life stat lines with variables attached to each statistic and set equal to a performance rating between 0 and 100. We then used a matrix to solve the system of equations for all six variables. The solution of the matrix gives us our coefficients for our formula for QPI.

In order to create our matrix, we need a series of data points to serve as baselines. In order to set the maximum rating, we selected some of the strongest quarterback performances within the last 20 years and set their value equal to 100. Likewise, we chose the worst quarterback performances over that same time period and set their value equal to 0. We chose this range in order to best reflect modern quarterback play, echoing our commitment to making QPI a rating system applicable to today's game. By setting up these stat lines in a matrix, we are able to use linear regression to find our coefficients. The matrix we used looks like this:

$$\begin{bmatrix} 0 & \frac{315}{27} & \frac{6}{27} & 0 & 0 & 0 \\ 0 & \frac{380}{24} & \frac{6}{24} & 0 & 0 & 0 \\ \frac{27}{9} & \frac{186}{19} & \frac{2}{19} & 0 & 1 & 2 \\ \frac{18}{5} & \frac{326}{28} & \frac{6}{28} & 0 & 0 & 0 \\ \frac{-2}{2} & \frac{462}{42} & \frac{7}{42} & 0 & 0 & 0 \\ \frac{181}{16} & \frac{263}{31} & \frac{2}{31} & \frac{1}{31} & 0 & 0 \\ \frac{4}{2} & \frac{66}{14} & 0 & \frac{5}{14} & 0 & 0 \\ 0 & \frac{73}{17} & 0 & \frac{4}{17} & 1 & 0 \\ 0 & \frac{35}{20} & 0 & \frac{4}{20} & 0 & 0 \\ \frac{2}{1} & \frac{23}{17} & 0 & \frac{1}{17} & 0 & 0 \\ 0 & \frac{212}{42} & \frac{1}{42} & \frac{7}{42} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2}$$

As the left matrix is not invertible, we must use a Moore-Penrose Pseudo-Inverse to solve for the coefficients. As this is a tall matrix (m $\times$ n with m $>$ n), the necessary formula to obtain the pseudo-inverse, A^+, of the matrix A with transpose A^T is:

$$A^+ = (A^T A)^{-1} A^T \tag{3}$$

By then multiplying that pseudo-inverse to the vector of assigned scores, we then solve for our coefficients. Our results for the coefficients using this method were:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0.39140497 \\ 5.382185724 \\ 189.9442196 \\ -80.91430327 \\ -7.423134039 \\ 18.45804009 \end{bmatrix} \tag{4}$$

Our overall QPI formula is as follows:

$$QPI = 0.3914(\frac{\text{rush yards}}{\text{attempt}}) + 5.3822(\frac{\text{pass yards}}{\text{attempt}}) + 189.9442(\frac{\text{passing touchdowns}}{\text{attempt}}) \\ -80.9143(\frac{\text{interceptions}}{\text{attempt}}) - 7.4231(\frac{\text{fumbles}}{\text{game}}) + 18.4580(\frac{\text{rushing touchdowns}}{\text{game}}) \tag{5}$$

After QPI is developed, the second step of the project is determining which quarterback has the best performance during "clutch" situations. The NFL records statistics during games specifically within seven points in the fourth quarter, so we used that data to define a "clutch" situation. We first implemented our QPI rating during these situations to get a rating for a quarterback during "crunchtime," then found the quarterback's overall rating over that same period of time. We compared the two, subtracting the overall rating from the rating specifically during crunchtime to create a quarterback's Clutch Factor (CF):

$$CF_g = QPI_c - \frac{QPI_1 + QPI_2 + QPI_3}{3} \tag{6}$$

This formula consists of CF_g, which is per game clutch factor, QPI_c, which is crunchtime QPI, and then QPI_1, QPI_2, and QPI_3, which are QPI for quarters one, two, and three respectively. We then combine all of a player's per-game CFs over the course of a season to create a seasonal CF value. Finally, we ranked all active quarterbacks based on their seasonal CF and compared with existing rankings.

4. Results

We tested our Clutch Factor formula on all thirty-two starting NFL quarterbacks of the 2023-24 season. Here are the results:

1. Desmond Ridder (32.6)	9. Baker Mayfield (6.9)	17. Dak Prescott (-4.1)	25. Patrick Mahomes (-15.2)
2. Kenny Pickett (15.4)	10. Mac Jones (6.9)	18. Gardner Minshew (-4.3)	26. Jordan Love (-15.7)
3. CJ Stroud (15.2)	11. Kyler Murray (6.5)	19. Kirk Cousins (-8.2)	27. Trevor Lawrence (-16.5)
4. Zach Wilson (13.9)	12. Justin Herbert (5.5)	20. Jalen Hurts (-9.1)	28. Lamar Jackson (-19.6)
5. Jared Goff (13.4)	13. Derek Carr (1.0)	21. Matt Stafford (-9.7)	29. Will Levis (-22.0)
6. Bryce Young (10.9)	14. Russell Wilson (0.3)	22. Joshua Dobbs (-13.2)	30. Brock Purdy (-26.9)
7. Joe Burrow (10.5)	15. Sam Howell (-0.7)	23. Josh Allen (-14.7)	31. Tua Tagovailoa (-32.3)
8. Jake Browning (9.1)	16. Geno Smith (-3.2)	24. Aidan O'Connell (-15)	32. Justin Fields (-36.5)

Figure 1. CF ranking of all 32 starting NFL quarterbacks of 2023

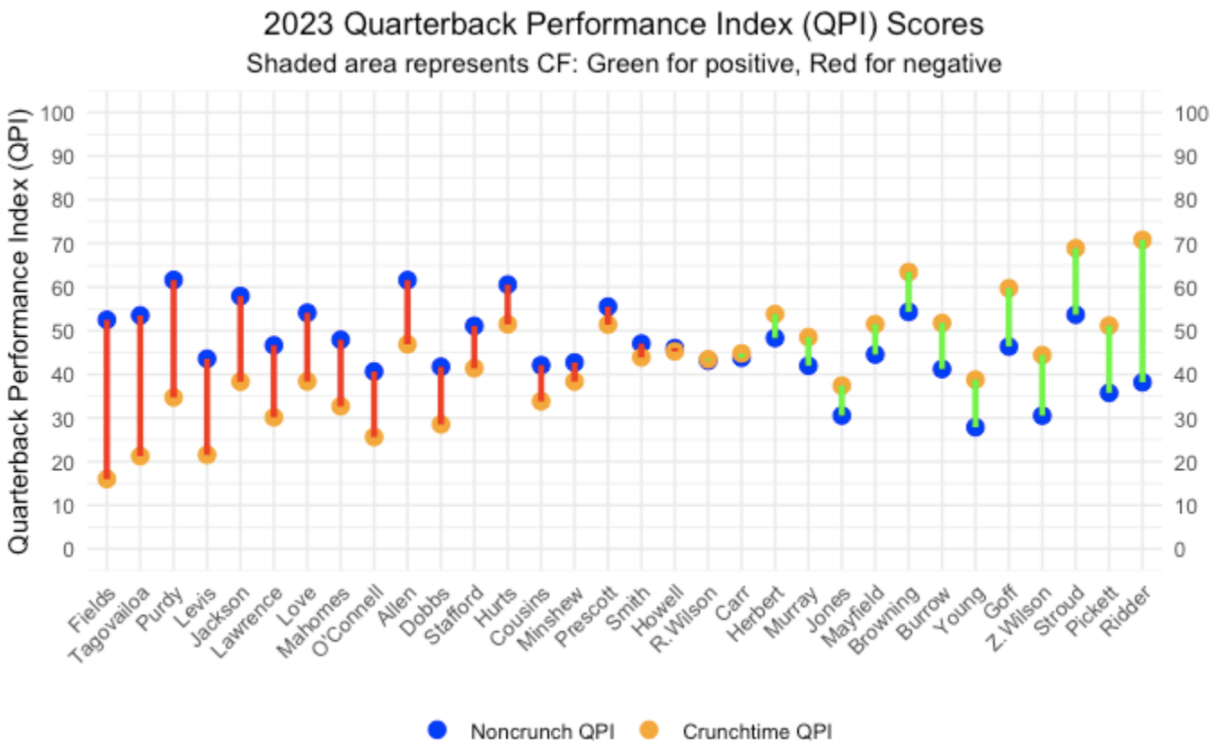

FIGURE 2.

Fourteen out of thirty-two starting quarterbacks in 2023 received positive CF scores, while eighteen received negative scores. Evidently, most quarterbacks actually played worse than their normal selves when the game was on the line. Whether this is a result of nerves, we cannot say for certain, but our ranking suggests that the majority of NFL quarterbacks in 2023 are not clutch.

As for individual players, unsurprisingly, our findings were far different from all other clutch rankings we have seen because of our unique definition of the word clutch. Patrick Mahomes, who ranked number one on CBS Sports' list[1], is ranked 24th on our list, having a Clutch Factor score of -15.2. Statistically, on average, he plays worse in the fourth quarter within seven points than he does in the first three quarters. C.J. Stroud, on the other hand, placed second in CBS Sports' ranking[1] and third in ours. He has a clutch factor of 15.2, proving that he statistically elevates his performance in crunchtime, when the game is on the line.

Desmond Ridder tops our ranking with an astounding CF score of 32.6, more than double our second place finisher, Kenny Pickett (15.4). As you can see in Figure 2, Ridder performed below average in the first three quarters on average, but became a completely improved version of himself in the fourth quarter within one score, having an impressive crunchtime QPI of 70.8.

Fascinatingly, many of the players with positive CF scores are generally considered subpar quarterbacks. As depicted in Figure 2, Desmond Ridder, Kenny Pickett, Zach Wilson, Bryce Young, and Mac Jones have the lowest five non-crunchtime QPI scores of any quarterback

that season. Then, in crunchtime, they increase their QPI scores by at least five. All five of these quarterbacks are young and inexperienced, so it is possible that it takes them more time than other quarterbacks to get settled into each game, which would explain the steep elevation in performance in the fourth quarter. It is also possible that these players are highly competitive, and performing better in the biggest moments is in their blood. Regardless of the reason, these quarterbacks deserve recognition for their late game improvement.

5. Discussion and future work

What sets our model apart from the rest is how we define and measure clutch ability. Clutch Factor takes into account typical performance stats for each quarterback in order to guarantee that we are answering the question of who is the most clutch, rather than who is the most talented. Clutch Factor also uses our up-to-date and improved metric for measuring quarterback performance (QPI). QPI properly weighs yards and touchdowns for both rushing and passing quarterbacks, while also appropriately penalizing players for making costly mistakes like throwing interceptions and fumbling the ball. Our matrix used to calculate our variable weights was created using relevant and accurate past stat lines. Our perfect games were crafted using the greatest games ever by the greatest quarterbacks of our generation, while our worst games were created using the most poor performances of all time.

While our model is accurate and successful in many ways, it falls short in a few areas. For one, it only has a few stat lines inputted into the matrix to produce the QPI coefficients. This hinders accuracy because it lacks a wide enough range of stats and scores to base its final numbers off of. The ratio of some of our coefficients could also be improved. For example, one passing touchdown is equivalent to throwing about 35 passing yards, which seems to be a slight inflation of the impact that passing yards truly have. Additionally, the model does not currently have a ceiling for a highest score a quarterback can receive. This could either be viewed positively or negatively. Some would argue that allowing QPI values greater than 100 makes understanding scores difficult because it strays away from the conventional understanding of 0-100 scoring. However, we believe that allowing very special performances to exceed 100 elevates the model because it always leaves room for players to improve their game, and hence, their QPI scores. As new all-time greats enter the game and the sport develops, we will likely see higher and higher QPIs.

The model could benefit in the future from transitioning to a more statistical approach. For example, in order to create accurate QPI numbers for the values in between 50 and 100, we could implement tools like quartiles and standard deviation. We also could improve the method of selecting all-time great quarterback games, by implementing some sort of statistical analysis beforehand. Instead of using a collection of rankings of best games from other sources, we could find some objective way to identify best games to eliminate any potential bias.

6. Conclusion

QPI and CF provide a modern way to properly assess a quarterbacks performance, both in "clutch" situations and overall. QPI addressed the inadequacies of the passer rating system traditionally used by the NFL, and creates a new objective measurement system. The roots of our coefficients used are easily traceable and produce results fairly consistent

with other metrics when tested. By including more relevant statistics and a clearly defined method, QPI is a new, updated version of passer rating. Our CF system takes a classic question and answers it in a different way, comparing quarterback "clutch" ability to their own average play, not against other quarterbacks in similar situations. Using this method, we have successfully taken overall talent out of the evaluation of whether a quarterback truly plays better in "crunchtime" situations. With this new system of measurement, we created a ranking of the most "clutch" quarterbacks which differs greatly from any prior ranking.

References

[1] D. Clawson, *Ranking NFL's top 10 clutch qbs: No surprise at No. 1, C.J. Stroud ranks high, Daniel Jones cracks list.* https://www.cbssports.com/nfl/news/ranking-nfls-top-10-clutch-qbs-no-surprise-at-no-1-c-j-stroud-ranks-high-daniel-jones-cracks-list/, July 2024. Date Accessed: October 3, 2024.

[2] A. Fromal, *NFL passer score: A new and improved statistic for measuring qb play.* https://bleacherreport.com/articles/553290-passer-score-a-new-and-improved-statistic-for-measuring-quarterback-play, December 2010. Date Accessed: October 3, 2024.

[3] M. Jetter, *Game, set, and match: Do women and men perform differently in competitive situations?*, Journal of Economic Behavior and Organization, 119 (2015), pp. 96–108.

[4] Rufio, *Why passer rating is not a good stat.* https://www.dawgsbynature.com/2013/12/4/5175398/why-passer-rating-is-not-a-good-stat, December 2013. Date Accessed: October 3, 2024.

[5] J. Schuhmann, *NBA storylines: Stephen Curry among top clutch performers.* https://www.nba.com/news/nba-storylines-2023-24-clutch-leaders, January 2024. Date Accessed: October 3, 2024.

Appendix A. Quarterback games Used to Develop QPI

The quarterbacks games we selected as our baseline values were as follows (with assigned scores):

$$\begin{bmatrix} \text{Aaron Rodgers 2014 Week 9} \\ \text{Tom Brady 2009 Week 6} \\ \text{Lamar Jackson 2023 Week 4} \\ \text{Patrick Mahomes 2018 Week 2} \\ \text{Peyton Manning 2013 Week 1} \\ \text{Colin Kaepernick 2012 Divisional Playoff} \\ \text{Nathan Peterman 2017 Week 10} \\ \text{Jake Delhomme 2009 Week 1} \\ \text{Peyton Manning 2015 Week 9} \\ \text{Derek Anderson 2009 Week 5} \\ \text{Ty Detmer 2001 Week 2} \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 100 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{7}$$

Memo (Letter)

Dear Representatives of the NFL Statistical Team,

We the members of Ice in the Pocket LLC have developed two new statistical models which will transform the way that you analyze a quarterback's performance. The Passer Rating system has been in use for over fifty years at this point, but can no longer be considered applicable to today's game. From a lack of information surrounding the basis of this rating to a complete omission of rushing performance, Passer Rating has a plethora of issues which must be addressed.

We would like to propose our new system, Quarterback Performance Index (QPI). Developed using linear algebra and using all-time great and poor quarterback performances as a basis, QPI provides a much needed revamp to the old, tired Passer Rating system. We found weights for each of six different quarterback statistics: Passing yards per attempt, passing touchdowns per attempt, rushing yards per attempt, rushing touchdowns per game, interceptions per attempt, and fumbles per game. Putting each weight together, you can plug in a quarterback performance over the course of a game, season, or even career to get an overall QPI rating. QPI can seamlessly supplant Passer Rating, as with an easily replicated method it provides full transparency for any score given. It also improves on the lack of foundation found in Passer Rating, as there is a clear basis for the creation of QPI's weights.

Our second statistical model answers another question at the forefront of any avid football fan's mind - who is the most clutch quarterback? We built off of QPI to develop a model that evaluates a quarterback's performance in both "crunchtime" and average play, to compare who truly excels in the clutch. By simply subtracting a quarterback's average QPI during the first three quarters from their performance in the fourth quarter within 7 points, we have created a new metric called Clutch Factor (CF).

CF addresses several issues with prior rankings of "clutchness," primarily by removing overall talent from the equation. A quarterback who performs well all game will naturally seem more clutch when they continue to play at a high level during a close game situation. However, this is not an accurate measurement of whether that quarterback actually improves their play in these situations. CF solves this issue by comparing a quarterback with himself, not with others in similar situations.

To exhibit the novelty of our model, we have sampled the top 32 quarterbacks from last season to rank their CFs. Our results show the stark inadequacies with other quarterback "clutch" ranking systems, with a nearly completely different top 10 scores. Quarterbacks such as Desmond Ridder or Kenny Pickett, traditionally not thought of as clutch, have the highest CF scores while star quarterbacks like Patrick Mahomes are towards the bottom.

Overall, QPI and CF both provide a breath of fresh air in the world of football statistical analysis as well as solve problems with traditional models. By implementing QPI and CF into NFL-sponsored publications and media, we can provide a new level of modernity and reliability in our quarterback analysis. Football lovers around the world will benefit with the NFL's adoption of QPI and CF. Thank you for your time and consideration.

Sincerely,
Ice in the Pocket LLC