Submitted to the Annals of Applied Statistics

RUSH INDEPENDENT PASSING PLAYER EFFICIENCY NUMBER (RIPPEN)

By Gregory J. Matthews

Skidmore College

AND

By Russell Cain

Loyola University Chicago

AND

By Donald Stolz

Smith College

RIPPEN, Rush Independent Passing Player Efficiency Number, is a new measurement of passer performance. In a simulated world, how would a passer perform starting from their twenty yard line and only performing pass plays? The aspects of each play are simulated using a Bayesian model. This allows rookies and backups with minimal data to be fairly evaluated. Drives would end in a touchdown, field goal or turnover. A player's RIPPEN is the average number of points they would be expected to score per drive. Our metric improves on existing passer rating systems because it is updated to current NFL data, does not weight passing touchdowns, and it is able to be more intuitively understood.

1. Introduction. The current passer rating measure has been around since 1973. NFL's Quarterback Rating:
Using the notation from van Dohlen (2011):

$$QBR = \left(\frac{\frac{C}{A} - 0.3}{0.2} + \frac{\frac{Y}{A} - 3}{4} + \frac{\frac{T}{A}}{0.05} + \frac{0.095 - \frac{I}{A}}{0.04} \right) \left(\frac{100}{6} \right)$$

where C = Number of Completions
Y = Number of Yards
A = Number of Attempts
T = Number of Touchdowns
I = Number of Interceptions

Each of these four components has a maximum of 2.375, so a "perfect" passer rating in the NFL is $\frac{(2.375)(4)(100)}{6} = 158.3$.
The NCAA passer rating is: $\frac{8.4Y+330T+100C-200I}{A}$

Keywords and phrases: sports analytics, Bayesian modeling, competitive balance, MCMC

This number ranged from -731.6 to 1261.6. Ridiculous.
Passer rating is bad. RIPPEN is better. The NCAA and NFL formulas are different. Mention this.

2. Articles. Tim Tebow example of why QBR is bad:

Read more about this (pareto-frontier). Might be interesting
- Would we add something like this to our results

DYAR and DVOA:

nih: charles poliquin

JQAS

A Statistical Analysis of NFL Quarterback Rating Variables
Derek Stimel, Journal of Quantitative Analysis in Sports

The Quarterback Prediction Problem: Forecasting the Performance of College Quarterbacks Selected in the NFL Draft
Julian Wolfson et al., Journal of Quantitative Analysis in Sports

Analyzing dependence matrices to investigate relationships between national football league combine event performances
Brook T. Russell et al., Journal of Quantitative Analysis in Sports

Isolating the Effect of Individual Linemen on the Passing Game in the National Football League
Benjamin C Alamar et al., Journal of Quantitative Analysis in Sports

Quantifying NFL Coaching: A Proof of New Growth Theory
Kevin P. Braig, Journal of Quantitative Analysis in Sports

CITE Passer Rating
CITE QBR

Don Steinberg: How I Learned to Stop Worrying and Love the Bomb

Quarterback Rating:

NFL Passer rating:

College Passer efficiency:

Defending Passer rating: Kerry Byrne

PRO FOOTBALL; The N.F.L.'s Passer Rating, Arcane and Misunderstood

Stimel (2009) Looking for structural breaks in QBR.
van Dohlen (2011)

2.1. *Criticism of QBR.* Arbitrary scale (0 to 158.3??) Hard to interpret (What does 121.6 mean?) QBR overly credits QBs for scoring TDs – discuss whether or not this is entirely wrong. Something to be said for "getting er done", but they weight this a bit too much for a metric which assesses QB efficacy.

3. Methods. We propose Rush Independent Passing Player Efficiency Number (RIPPEN). Describe what we did.
Loose Outline

3.1. *Data - Open Source.* RIPPEN uses the exact same inputs as QBR: Completions, Yards, Interceptions, and Touchdowns.

3.1.1. *nflscrapr.* Link in Rob Yurko - maybe describe process of gathering the data? Probably a good chance to introduce the few variables we actually look for and build with.

3.1.2. *our package (check it out!!!).* Exactly! Great time to discuss the variables we use.

3.2. *How we use our data.*

3.2.1. *Simulation! Bayseian?.* I think we have a prior, look at probability within, run it millions of times.. something like this.

3.2.2. *Markov Chain Notion:.* I know what this is! Describe overall concept and explain how it applies to what we are doing – borderline survival analysis.

-Markov-	Down 1	Down 2	Down 3	Down 4
Down 1	a	b	0	0
Down 2	c	0	d	0
Down 3	e	0	0	f
Down 4	0	0	0	1

$$
\begin{aligned}
a &= Pr(y_{d,1} > 10) \\
b &= 1 - a \\
c &= Pr(y_{d,2} > 10 - y_{d,1}) \\
d &= 1 - c \\
e &= Pr(y_{d,3} > 10 - y_{d,2} - y_{d,1}) \\
f &= 1 - e
\end{aligned}
$$

3.2.3. *Variable description! (More i's than Mississippi).* Dedicated to the variables noted in Markov Chain – make sense of each one and explain significance.

1. $y_{d,i}$: The i^{th} pass of the d^{th} down series. Therefore, $i \in 1, 2, 3, 4$ and d is loosely less than 8.

2. G: The result of the drive/simulation. Either 7 for TD, 3 for FG or 0 for interception or missed FG.

3. I(...): Indicator function: ...

4. $C_{d,i}$ ...

5. $I_{d,i}$: E[I(D = 4)] = P(D = 4)

6. $C_{d,i}$: $t_1' \cdot M = t_2' = [a\ b\ 0\ 0]$

7. .... : $t_2' \cdot M = t_3' = [a^2 + bc\ ab\ bd\ 0]$

8. $$Pr(G_j = 3) = Pr(FG \cap (\sum_{i=1}^{n-1} I(D_n = 4) = 0) \cap (\sum_{i=1}^{n-1} y_i < 80 | Q = \sum_{i=1}^{n-1} y_i)) \cdot P(Q = q)$$

   ... Pr(FG ∩ Q = q)

9. $$Pr(G = 7) = \sum_{n=1}^{\infty} Pr(\sum_{i=1}^{n} y_i > 80 | \sum_{i=1}^{n} I(D_i = 4) = 0) \cdot P(...)$$

3.3. *How we visualize, parse our analyses?.* Idk, Look at other sections of this paper and prep for that. Suppose we could at least speak to breaking it down by season, game, player and whatnot.

3.4. *Theoretical Results.* Do we have any?

3.5. *Correlation between RIPPEN and winning.* Compare RIPPEN and winning to QBR and winning.

3.6. *Preliminary Results & Notes.*

4. Results.

4.1. *Bayesian Posterior Distributions Stuff.* What do the posterior parameters look like?

4.2. *Rodgers vs Tebow Example. .*

4.3. *Distribtuion of RIPPEN.*

4.4. *Best Games/Seasons.*

5. Conclusion and Future Work. RIPPEN is good. We will do more eventually.

Adding a defensive adjustment.

Do we even want to add these things? How do we deal with pass interference? Defensive Holding? Sacks? Add another layer. Fumbles? Could treat similar to interceptions? Should interceptions ever result in negative numbers? How do we assign the negative numbers for interceptions?

References.

STIMEL, D. (2009). A Statistical Analysis of NFL Quarterback Rating Variables. *Journal of Quantitiative Analysis in Sports* **5** 1.

VAN DOHLEN, P. (2011). Tweaking the NFL's Quarterback Passer Rating for Better Results. *Journal of Quantitiative Analysis in Sports* **7** 22.

E-MAIL: gmatthews1@luc.edu E-MAIL: rcain@luc.edu

E-MAIL: dstolz@luc.edu