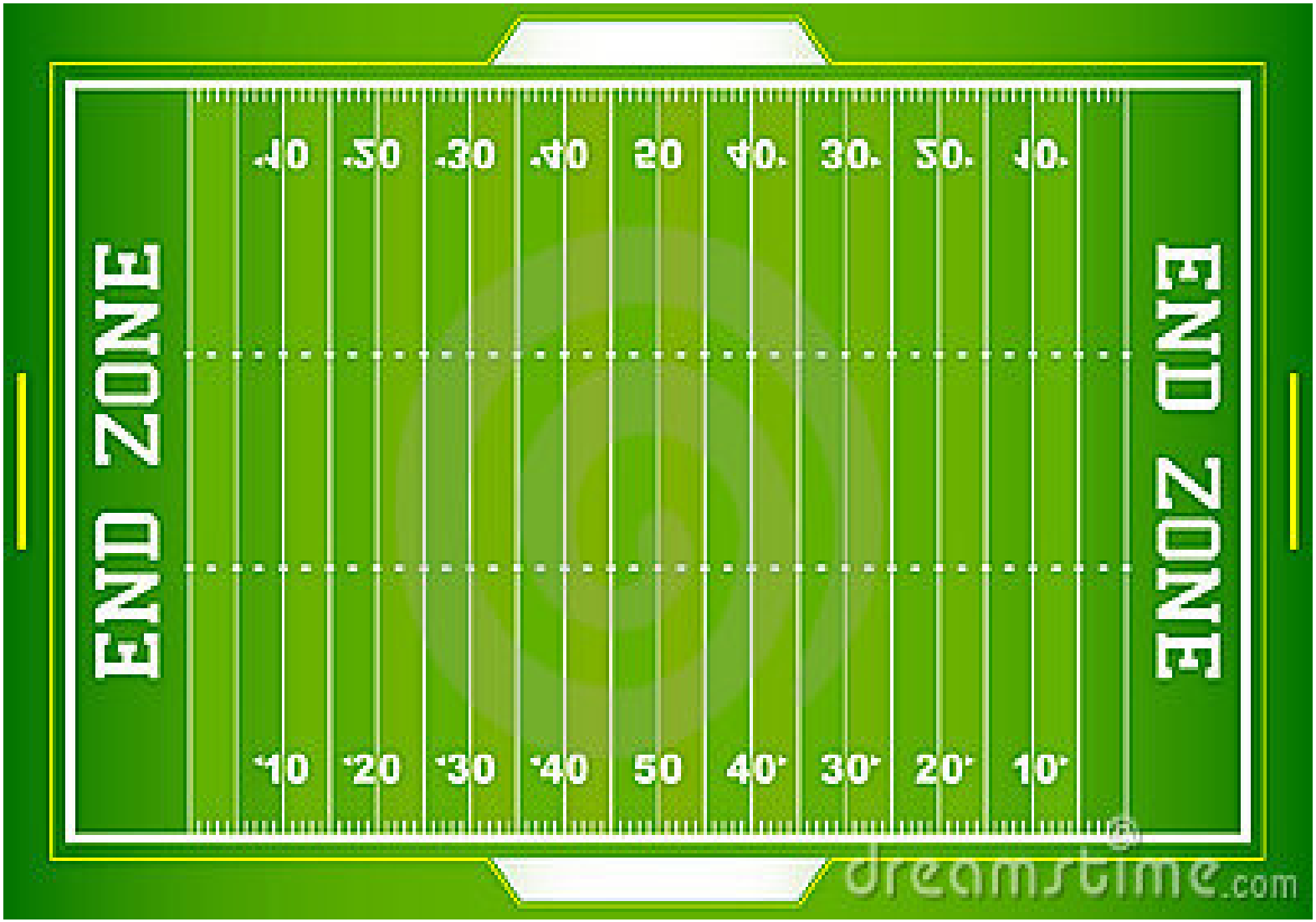
NFL Player Evaluation Using Expected Points Added with `nflscrapR`

Ron Yurko Sam Ventura Max Horowitz

Department of Statistics
Carnegie Mellon University

Great Lakes Analytics in Sports Conference, 2017

Game of Yards?



Tale of Two Runs

Arizona Cardinals line up on 3rd-and-20 on their own 10 yard line:

- David Johnson gains 11 yards, team punts next play

Tale of Two Runs

Arizona Cardinals line up on 3rd-and-20 on their own 10 yard line:

- David Johnson gains 11 yards, team punts next play

Pittsburgh Steelers line up on 3rd-and-1 on their own 10 yard line:

- Le'Veon Bell gains 4 yards, converts first down

Tale of Two Runs

Arizona Cardinals line up on 3rd-and-20 on their own 10 yard line:

- David Johnson gains 11 yards, team punts next play

Pittsburgh Steelers line up on 3rd-and-1 on their own 10 yard line:

- Le'Veon Bell gains 4 yards, converts first down

Johnson gained nearly 3 times more yards than Bell...

Tale of Two Runs

Arizona Cardinals line up on 3rd-and-20 on their own 10 yard line:

- David Johnson gains 11 yards, team punts next play

Pittsburgh Steelers line up on 3rd-and-1 on their own 10 yard line:

- Le'Veon Bell gains 4 yards, converts first down

Johnson gained nearly 3 times more yards than Bell...

NOT ALL YARDS ARE CREATED EQUAL

How to Value a Play? Enter the Statistician

How do we properly evaluate a play?

3rd-and-20, CLE 10, (10:48) (Shotgun) D.Johnson up the middle to CLE 21 for 11 yards (R.McLeod).

How to Value a Play? Enter the Statistician

How do we properly evaluate a play?
3rd-and-20, CLE 10, (10:48) (Shotgun) D.Johnson up the middle to CLE 21 for 11 yards (R.McLeod).

Consider a **play-by-play** dataset:

- **Down:** 4 downs to advance the ball 10 (or more) yards
- **Yards to go:** distance in yards to convert first down
- **Yard line:** distance in yards away from opponent's endzone (100 to 0) - the field position
- **Time remaining:** seconds remaining in half, each half is 1800 seconds long (overtime is 900)

How to Value a Play? Enter the Statistician

How do we properly evaluate a play?
3rd-and-20, CLE 10, (10:48) (Shotgun) D.Johnson up the middle to CLE 21 for 11 yards (R.McLeod).

Consider a **play-by-play** dataset:

- **Down:** 4 downs to advance the ball 10 (or more) yards
- **Yards to go:** distance in yards to convert first down
- **Yard line:** distance in yards away from opponent's endzone (100 to 0) - the field position
- **Time remaining:** seconds remaining in half, each half is 1800 seconds long (overtime is 900)

Value of a play depends on the situation

We Know This Already

Operations Research on Football (Carter and Machol, 1971)

- Former BYU and NFL quarterback Virgil Carter
- Took 2852 1st-and-10 plays, turned the field into 10 yard buckets, **averaged the value of the next scoring event**

We Know This Already

Operations Research on Football (Carter and Machol, 1971)

- Former BYU and NFL quarterback Virgil Carter
- Took 2852 1st-and-10 plays, turned the field into 10 yard buckets, **averaged the value of the next scoring event**

The Hidden Game of Football (Carroll et al., 1988)

- Classic work and first deep dive into football statistics
- Play's success is a function of down and yards to go
- Linear **expected points** model from -2 on team's goal line to +6 on opponent's, every 25 yards leads to 2 more expected points

We Know This Already

Operations Research on Football (Carter and Machol, 1971)

- Former BYU and NFL quarterback Virgil Carter
- Took 2852 1st-and-10 plays, turned the field into 10 yard buckets, **averaged the value of the next scoring event**

The Hidden Game of Football (Carroll et al., 1988)

- Classic work and first deep dive into football statistics
- Play's success is a function of down and yards to go
- Linear **expected points** model from -2 on team's goal line to +6 on opponent's, every 25 yards leads to 2 more expected points

Recent developments by Aaron Schatz at Football Outsiders, Brian Burke at ESPN, and Keith Goldner at numberFire

Reproducible with `nflscrapR`

Recent work in football analytics is not easily reproducible:

- Reliance on proprietary and costly data sources
- Data quality relies on potentially biased human judgement

Reproducible with nflscrapR

Recent work in football analytics is not easily reproducible:

- Reliance on proprietary and costly data sources
- Data quality relies on potentially biased human judgement

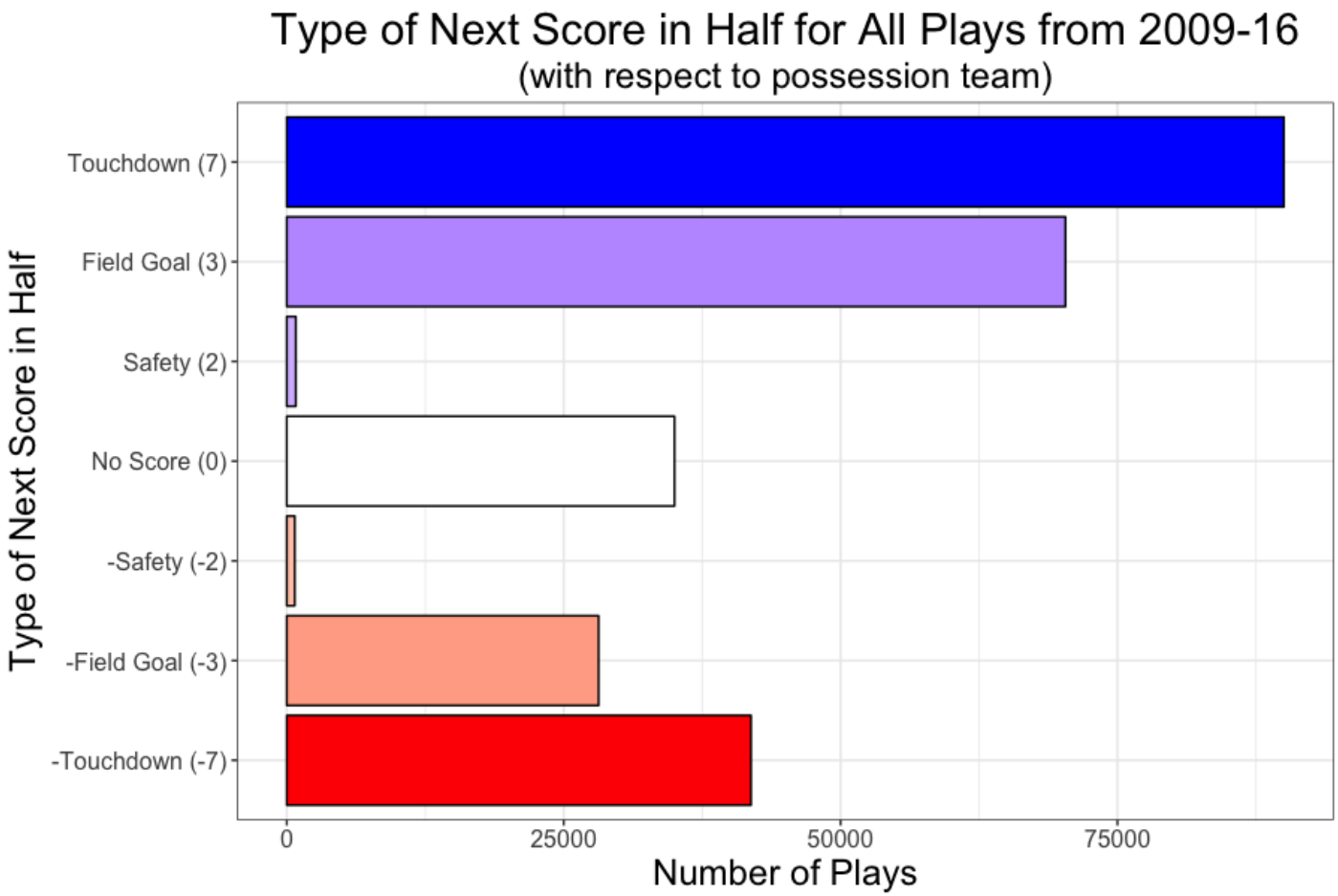
nflscrapR:

- R package created by Maksim Horowitz to enable easy data access and promote reproducible NFL research
- Collects play-by-play data from NFL.com and formats into R data frames
- Data is available for all games starting in 2009

Available on Github, install with:

```
devtools::install_github(repo=maksimhorowitz/nflscrapR)
```

How to Model Expected Points



Developed a novel **multinomial logistic regression** model built on `nflscrapR` data from 2009-2016 seasons, using the `nnet` package

Expected Points Model

Model is generating probabilities, agnostic of value associated with each next score type

Next Score: $Y \in \{$Touchdown (7), Field Goal (3), Safety (2), No Score (0), -Safety (-2), -Field Goal (-3), -Touchdown (-7)$\}$

Situation: $X \in \{$down, yards to go, field position, time remaining$\}$

Outcome probabilities: $P(Y = y|X)$

Expected Points (EP) $= E(Y|X) = \sum_y P(Y = y|X) * y$

Expected Points Added (EPA)

Expected Points Added (EPA) estimates a play's value based on the change in situation, providing a point value

- $EPA_{play_i} = EP_{play_{i+1}} - EP_{play_i}$

Expected Points Added (EPA)

Expected Points Added (EPA) estimates a play's value based on the change in situation, providing a point value

- $EPA_{play_i} = EP_{play_{i+1}} - EP_{play_i}$

David Johnson's 11 yards leading to 4th down?

Expected Points Added (EPA)

Expected Points Added (EPA) estimates a play's value based on the change in situation, providing a point value

- $EPA_{play_i} = EP_{play_{i+1}} - EP_{play_i}$

David Johnson's 11 yards leading to 4th down? 0.0096 EPA

Expected Points Added (EPA)

Expected Points Added (EPA) estimates a play's value based on the change in situation, providing a point value

- $EPA_{play_i} = EP_{play_{i+1}} - EP_{play_i}$

David Johnson's 11 yards leading to 4th down? 0.0096 EPA

Le'Veon Bell's 4 yards converting 1st down?

Expected Points Added (EPA)

Expected Points Added (EPA) estimates a play's value based on the change in situation, providing a point value

- $EPA_{play_i} = EP_{play_{i+1}} - EP_{play_i}$

David Johnson's 11 yards leading to 4th down? 0.0096 EPA

Le'Veon Bell's 4 yards converting 1st down? 0.9225 EPA

His 4 yards are almost 100 times more valuable!

EPA-based Metrics

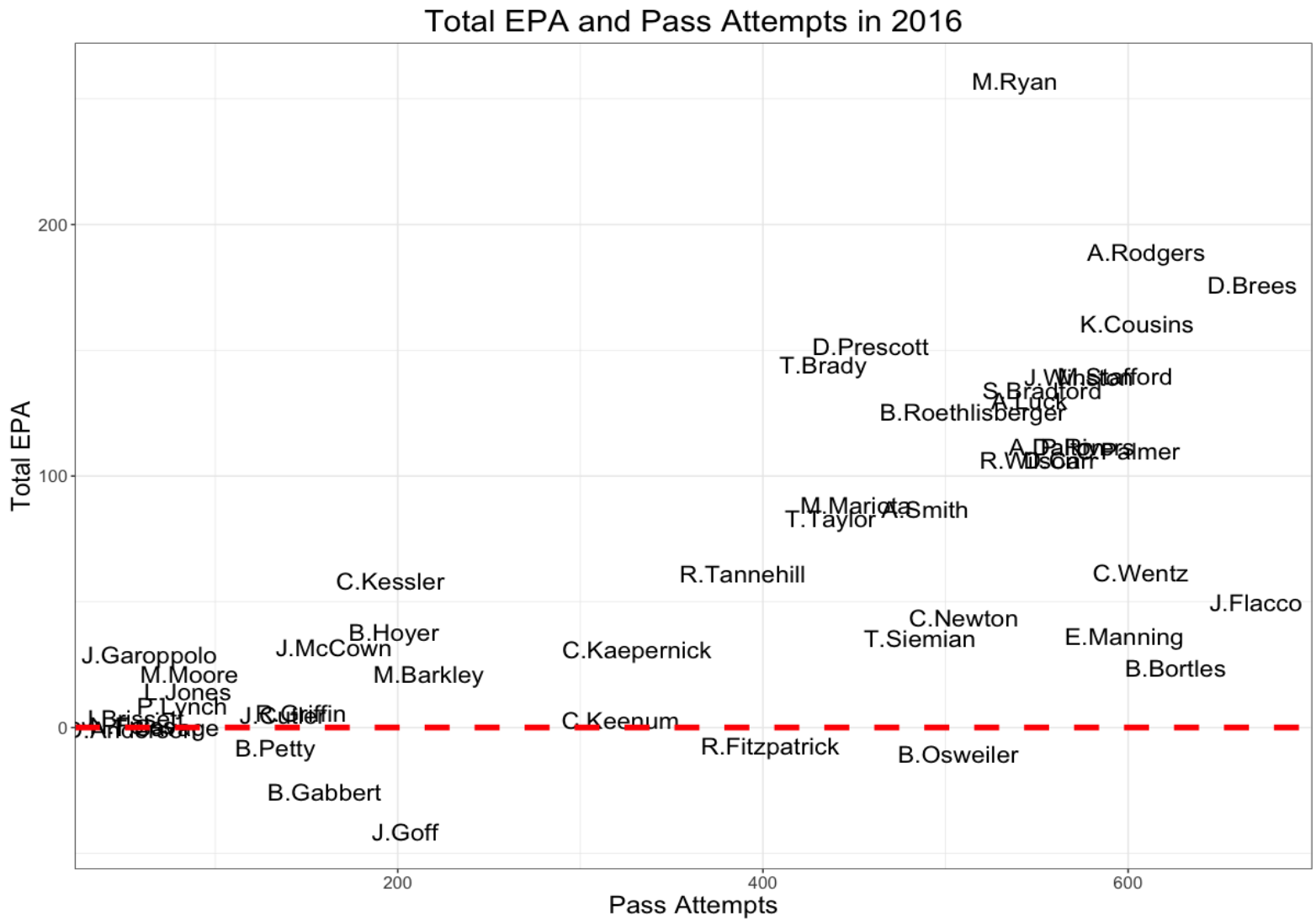
Using EPA, can calculate new metrics to evaluate players

The following slides will go through various EPA-based metrics:

- Total EPA
- Dividing by attempts for rate measure
- Introduce new weighted metrics

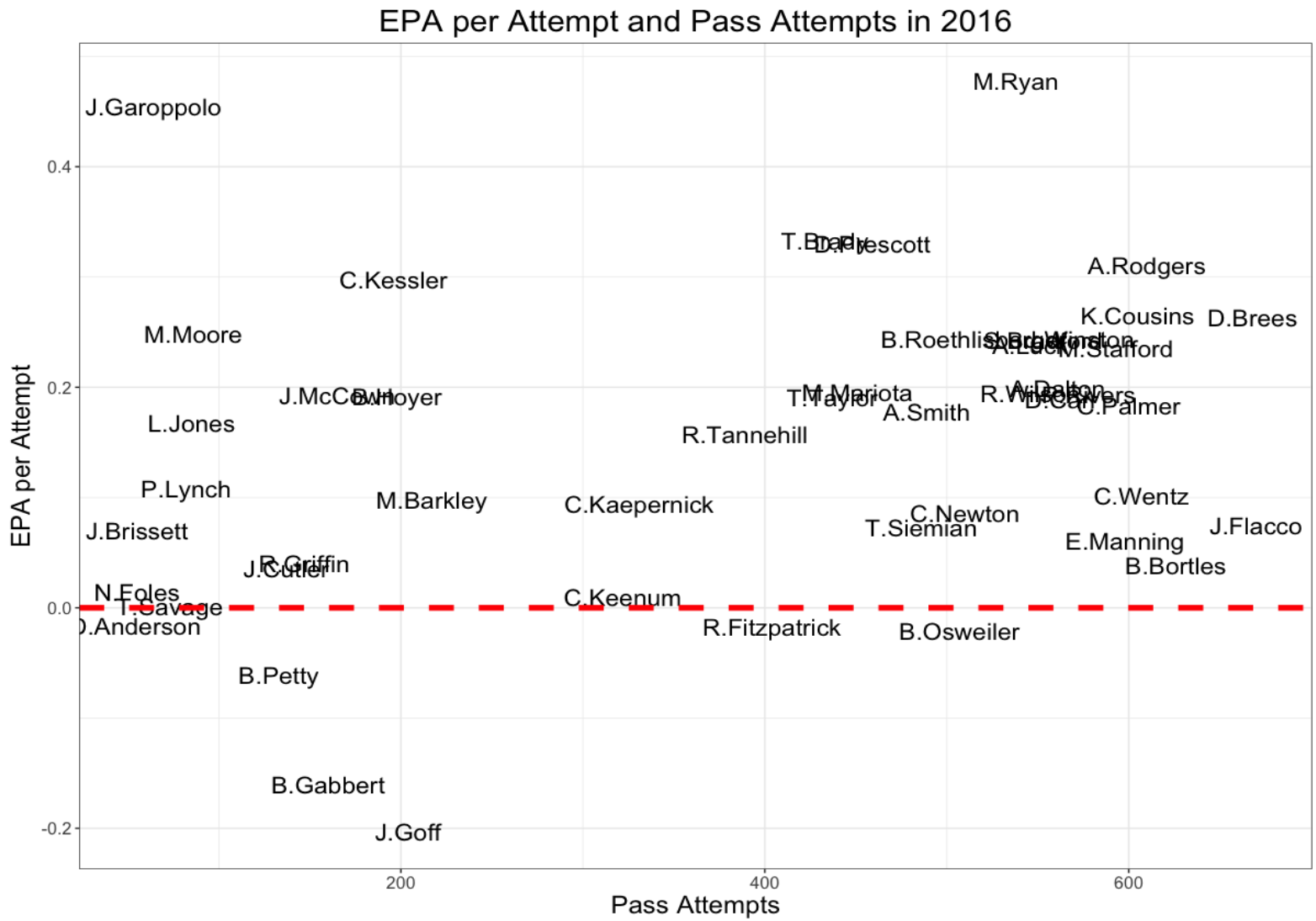
Quarterbacks: Total EPA

Total EPA $= \sum^{Attempts} EPA$



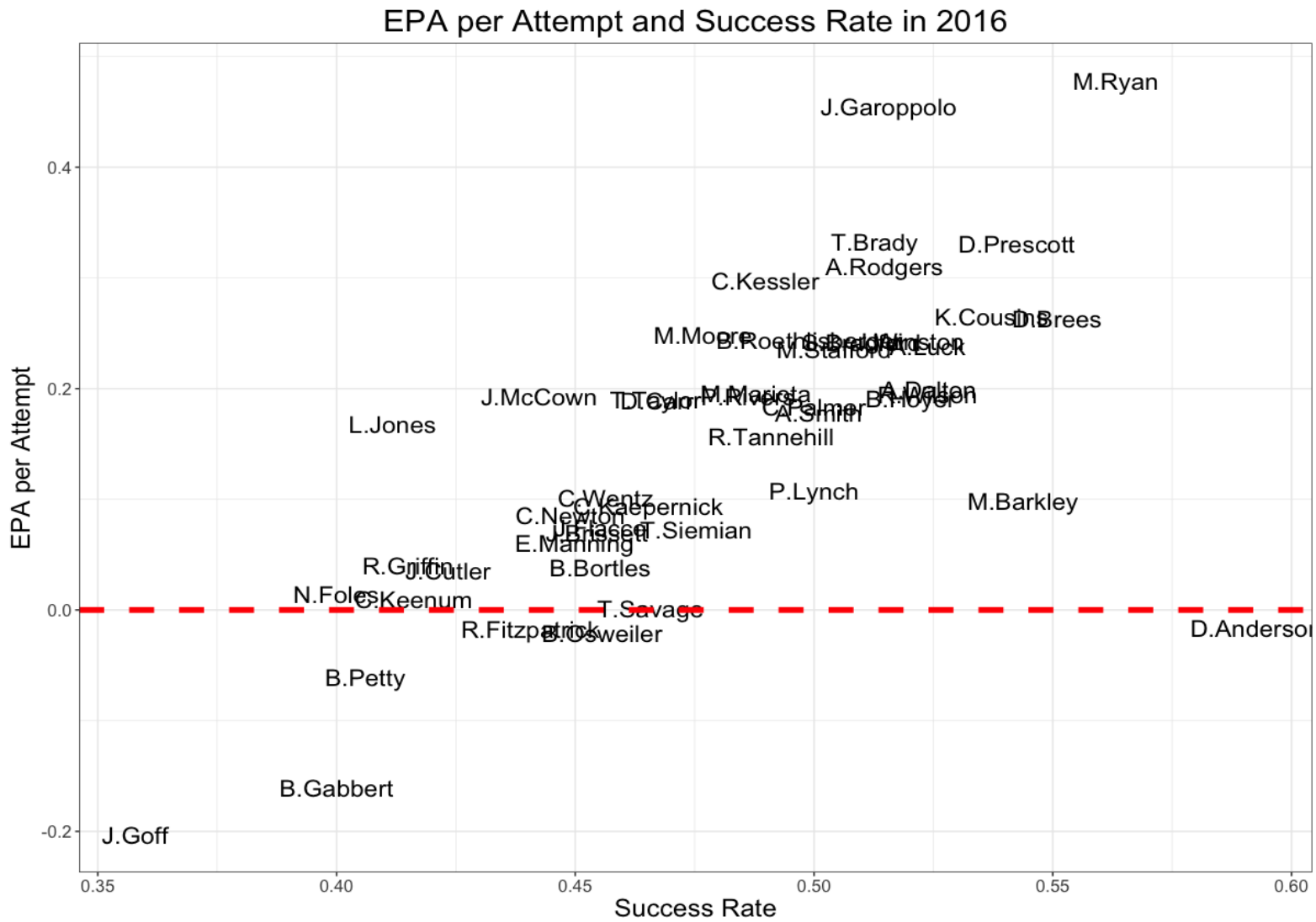
Quarterbacks: EPA per Attempt

EPA per Attempt $= \frac{Total\ \ EPA}{\#\ \ of\ \ Attempts}$



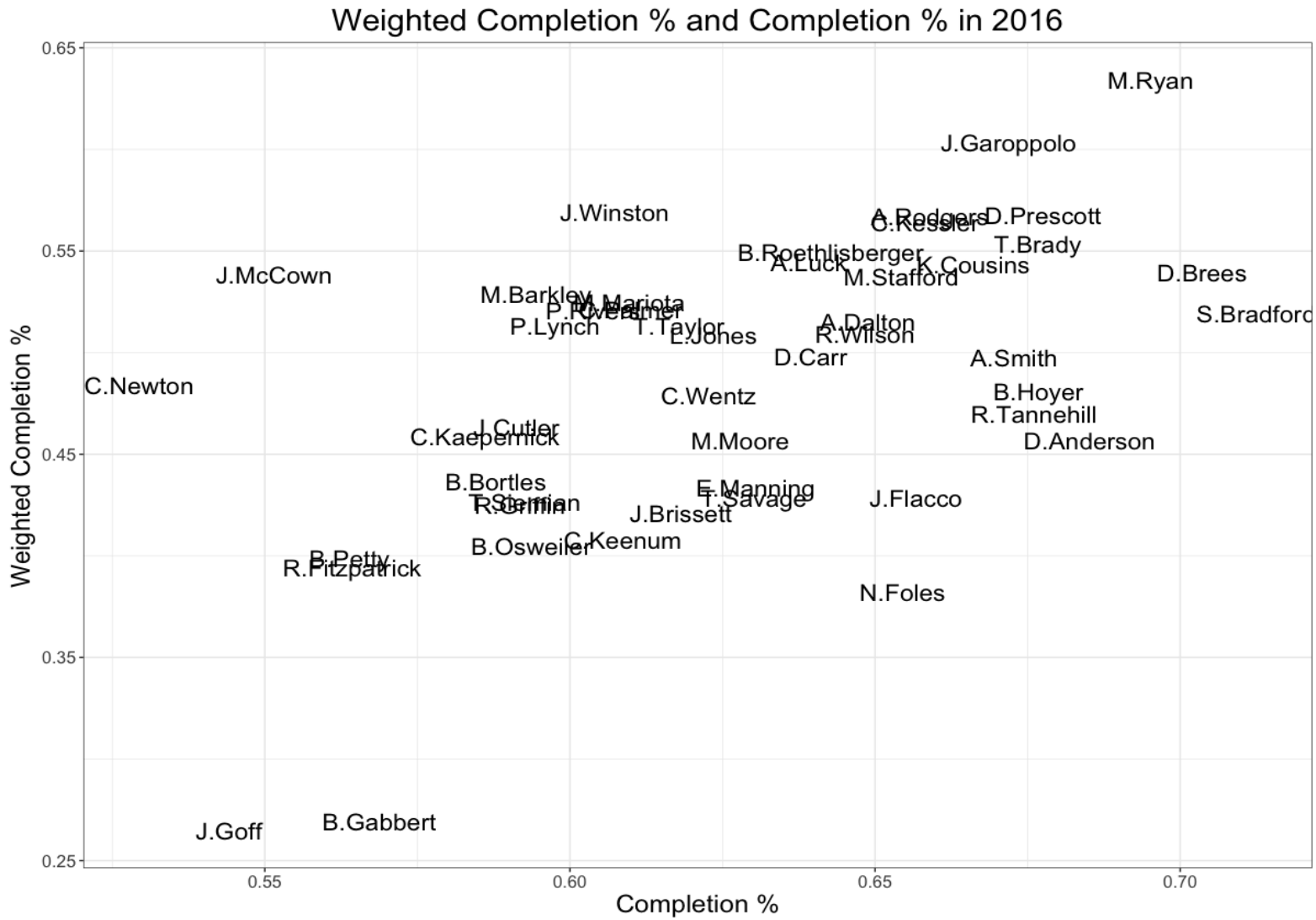
Quarterbacks: Success Rate

Success Rate $= \frac{\#\ plays\ with\ EPA>0}{\#\ plays}$ *(Berri & Burke, 2012)*



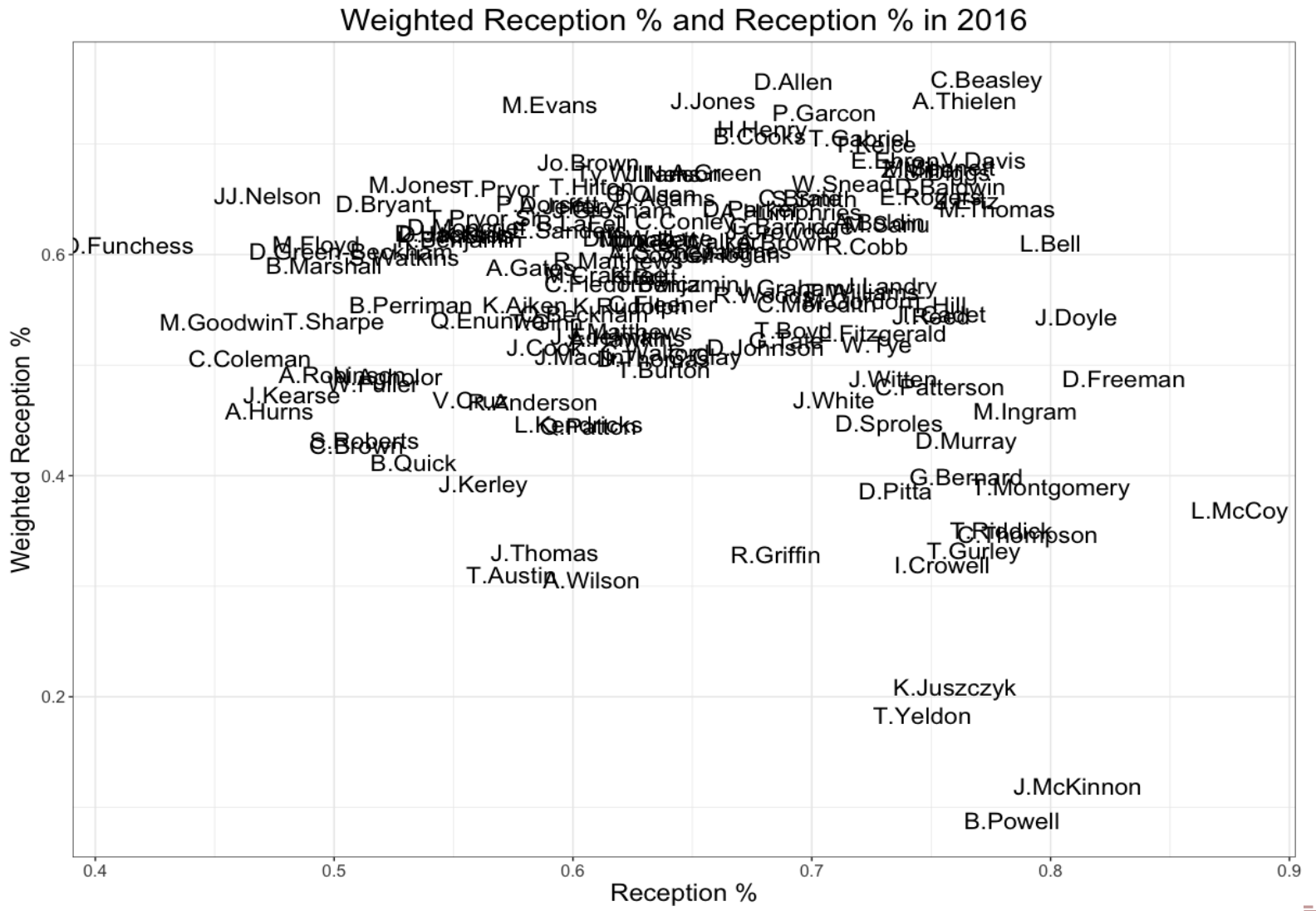
Quarterbacks: Weighted Completion %

$$\textbf{Weighted Completion \%} = \frac{\sum^{Completions} EPA}{\sum^{Attempts} |EPA|}$$



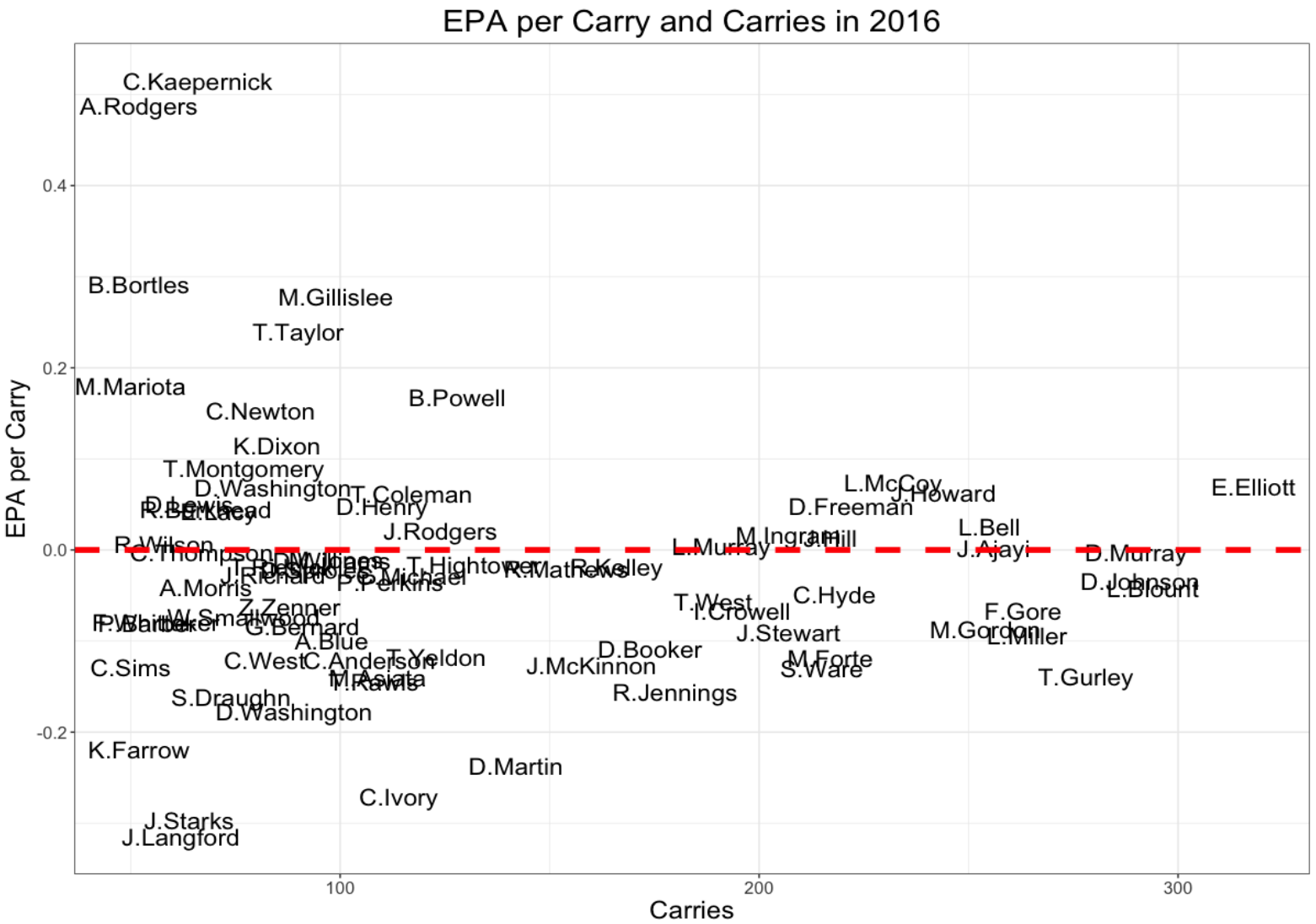
Receivers: Weighted Reception %

$$\textbf{Weighted Reception \%} = \frac{\sum^{Receptions} EPA}{\sum^{Targets} |EPA|}$$



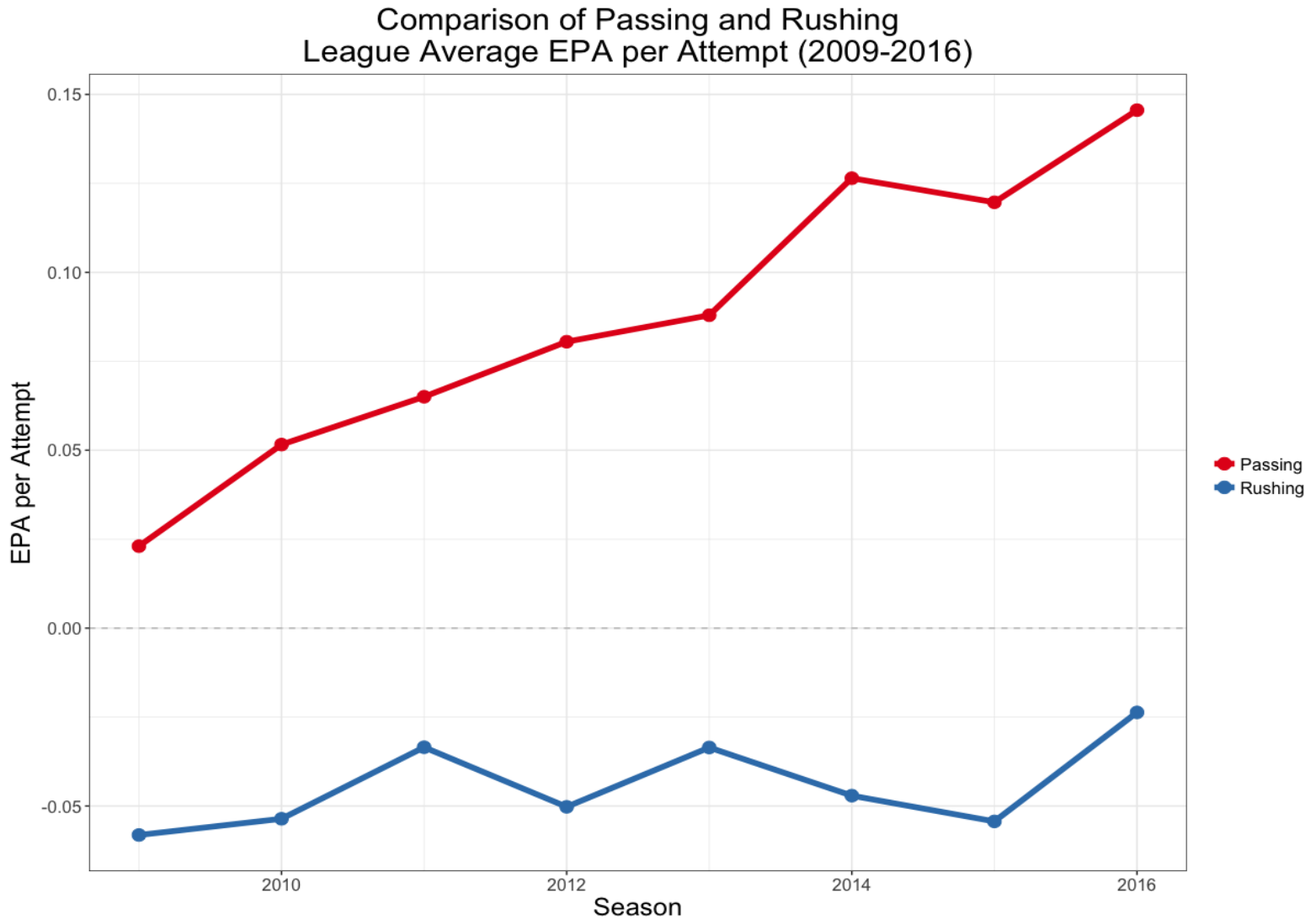
Rushers: EPA per Carry

EPA per Carry $= \frac{Total\ \ EPA}{\#\ \ of\ \ Carries}$

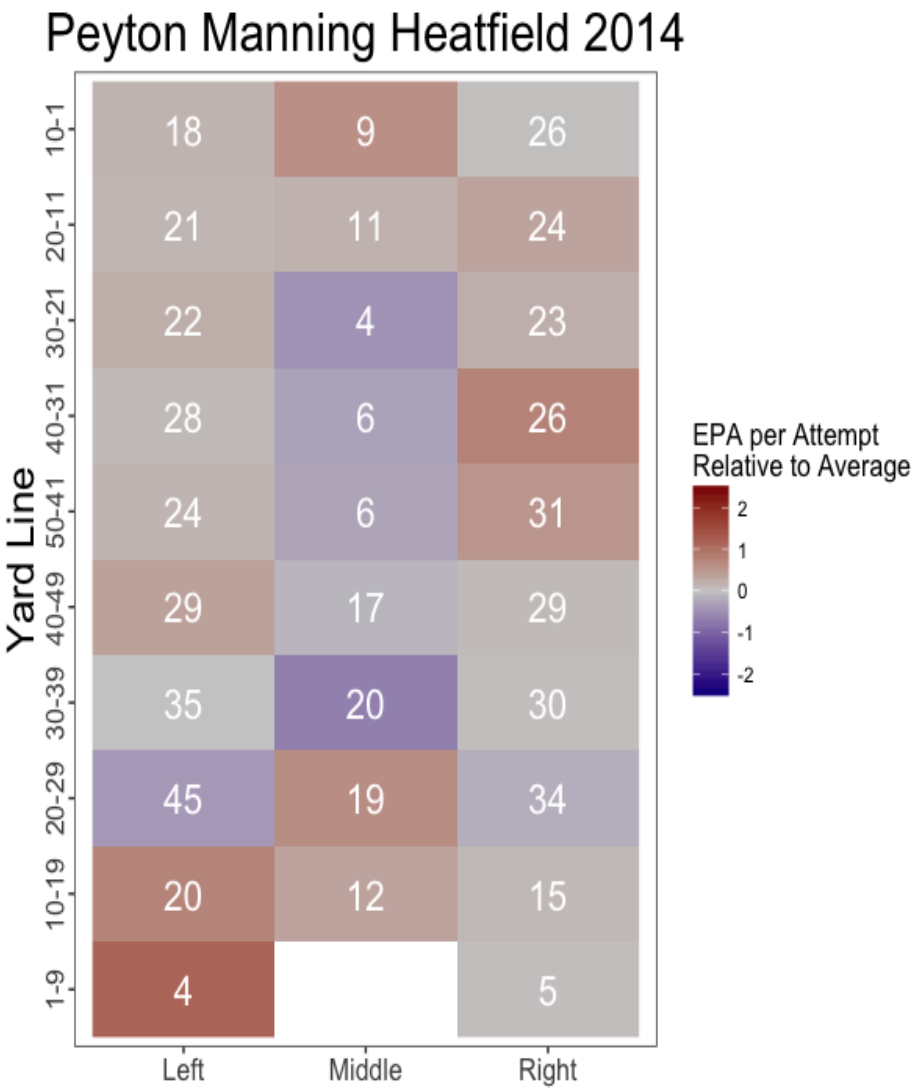


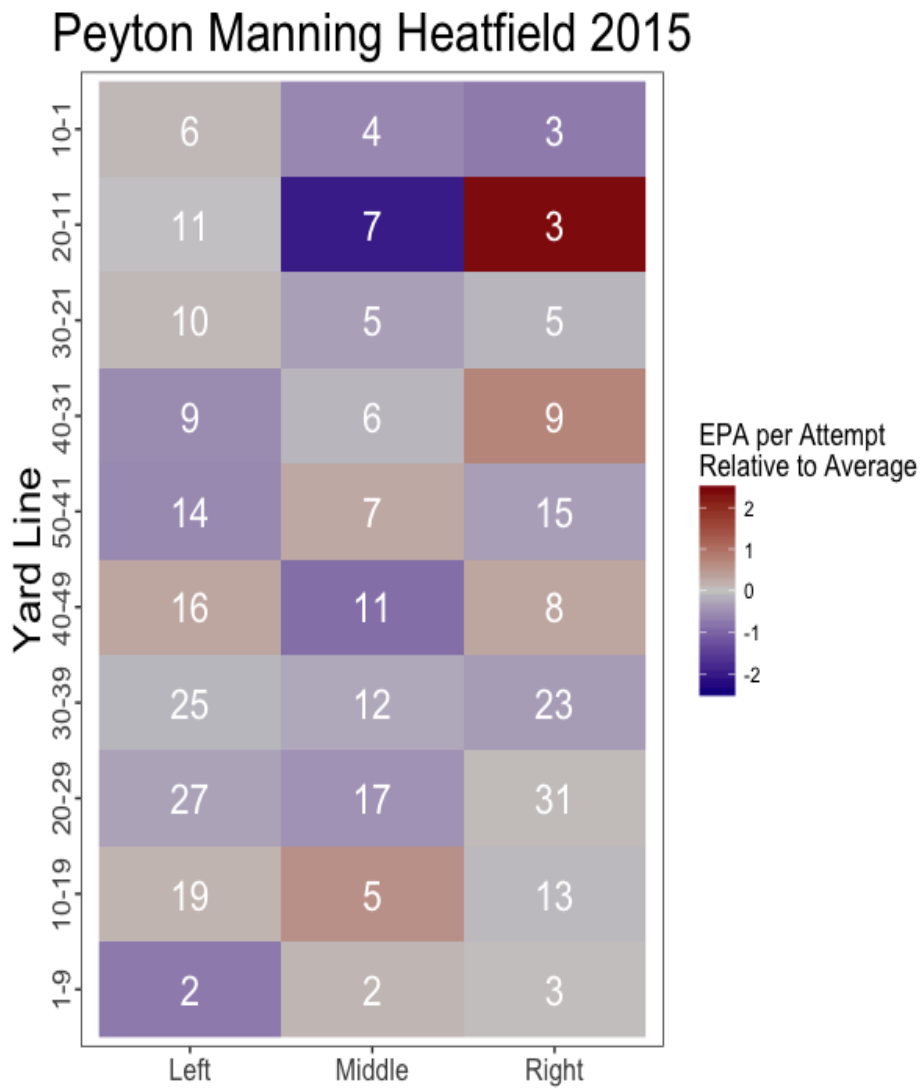
Difference Between Passing and Rushing

Rushing is costly, need to evaluate players relative to average

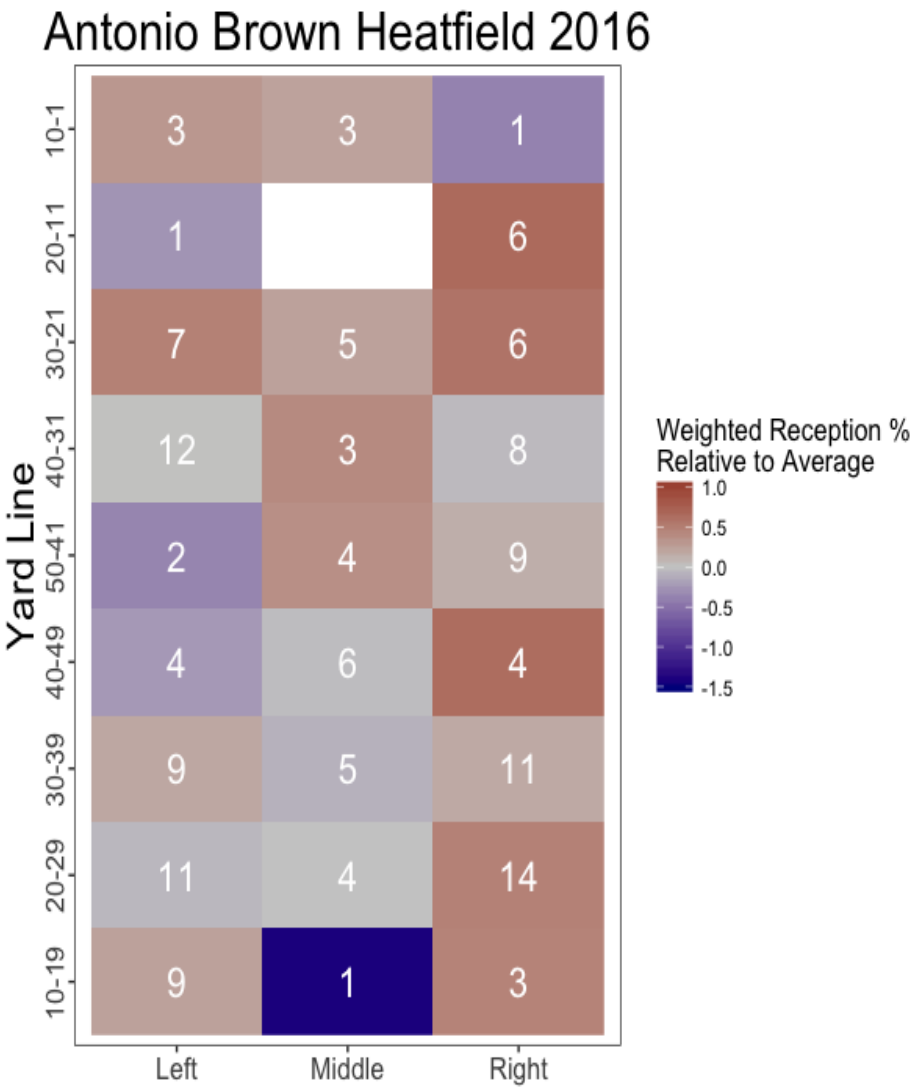


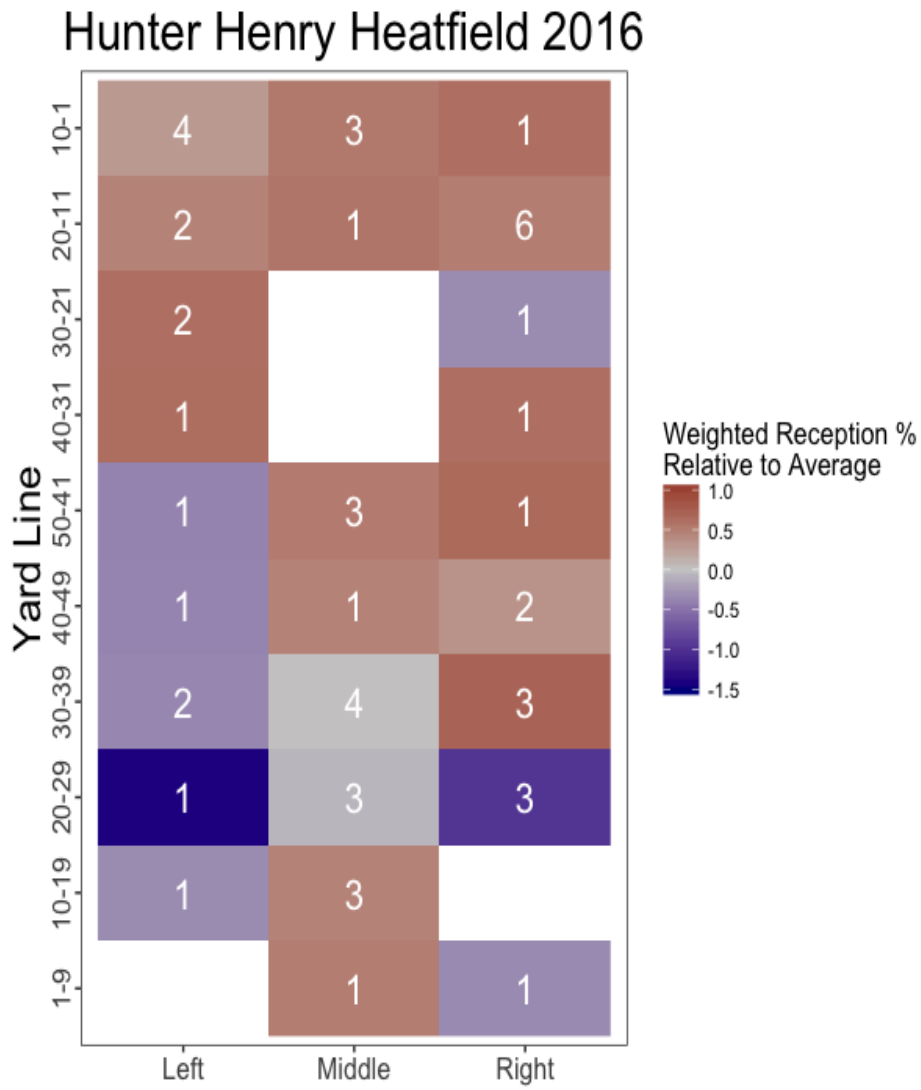
Peyton Manning Heatfield Comparison





Receiving Heatfield





Stabilization

Important to understand when stats stabilize to "trust" them

- *e.g. How many pass attempts does Dak Prescott need before we can believe the results?*

Stabilization

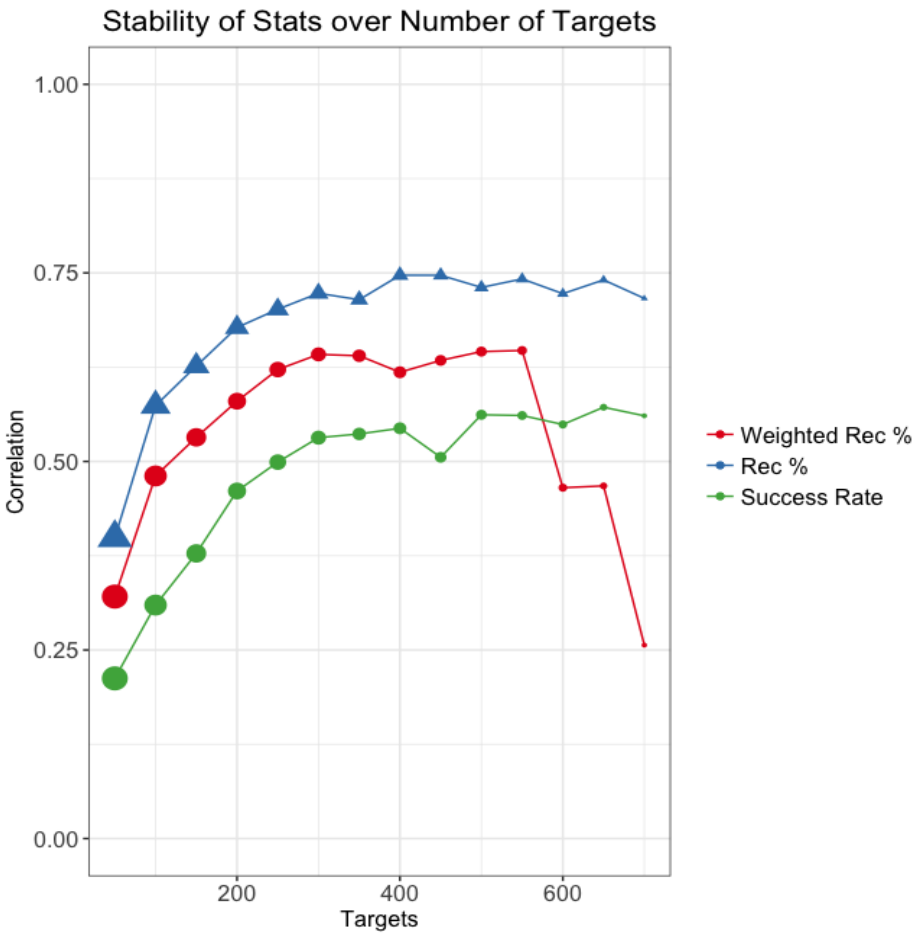
Important to understand when stats stabilize to "trust" them

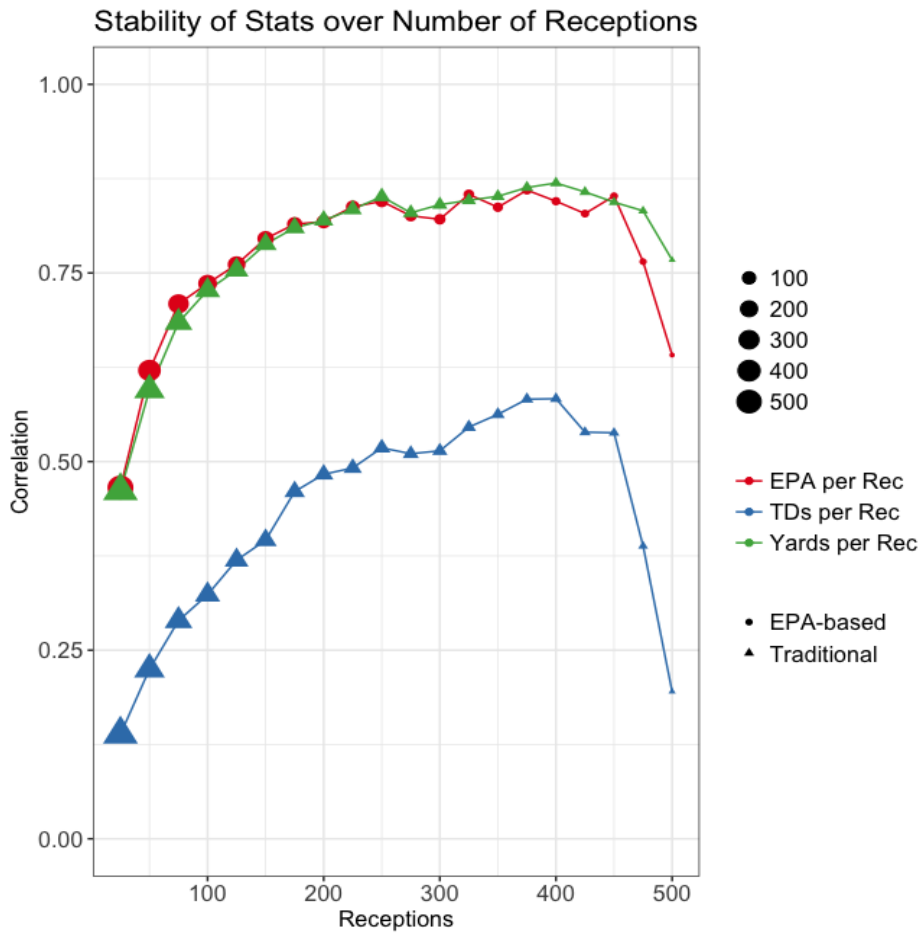
- *e.g. How many pass attempts does Dak Prescott need before we can believe the results?*

Follow approach by Tom Tango for baseball stats:

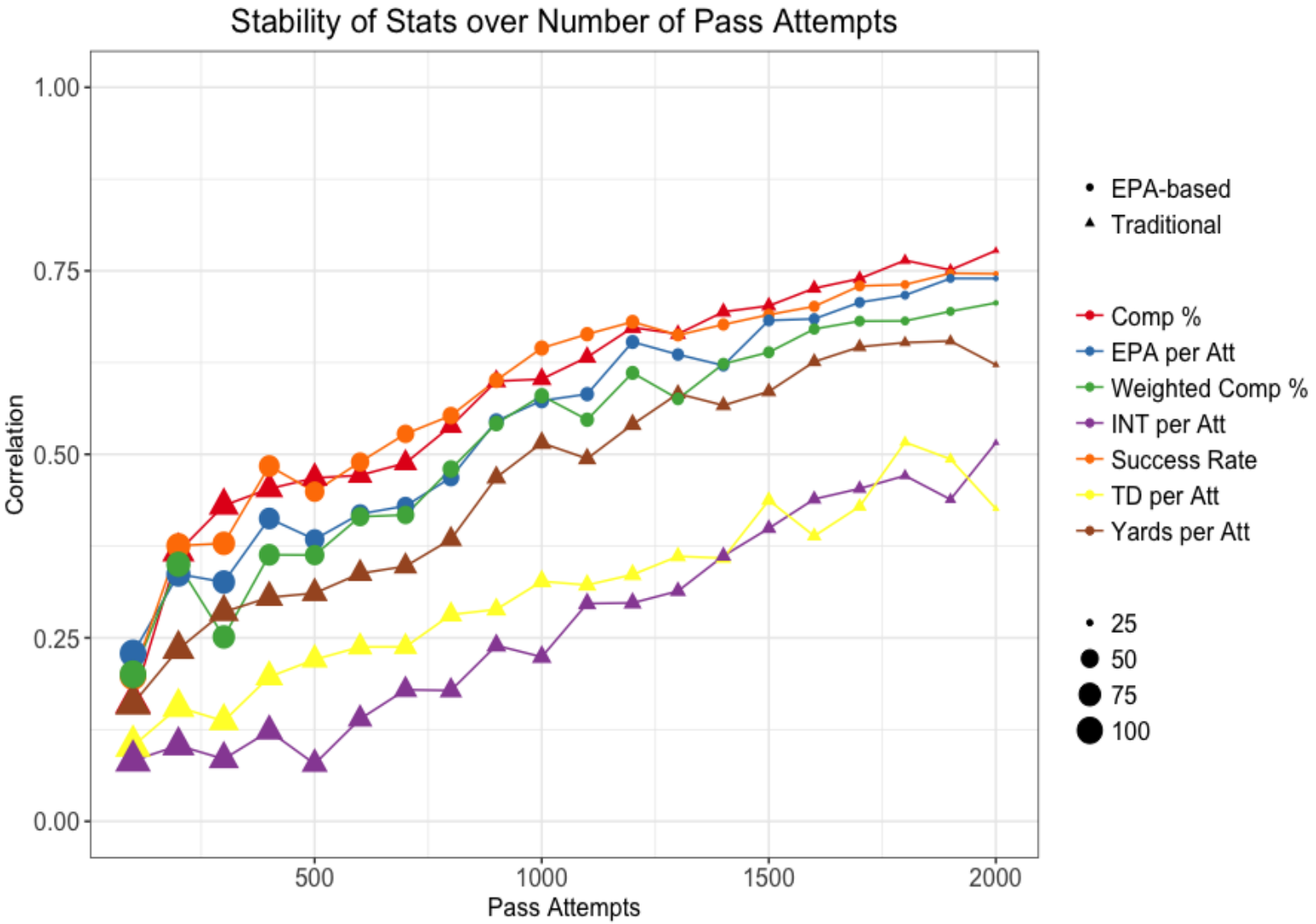
- For given number N of attempts (or targets, carries, etc.), find the players with at least that number
- For each player, randomly select two samples of size $\frac{N}{2}$ and calculate the stat of interest, repeat
- Compute the correlation r and repeat the process for each considered N, identify when correlation stabilizes

Receiving Stabilization

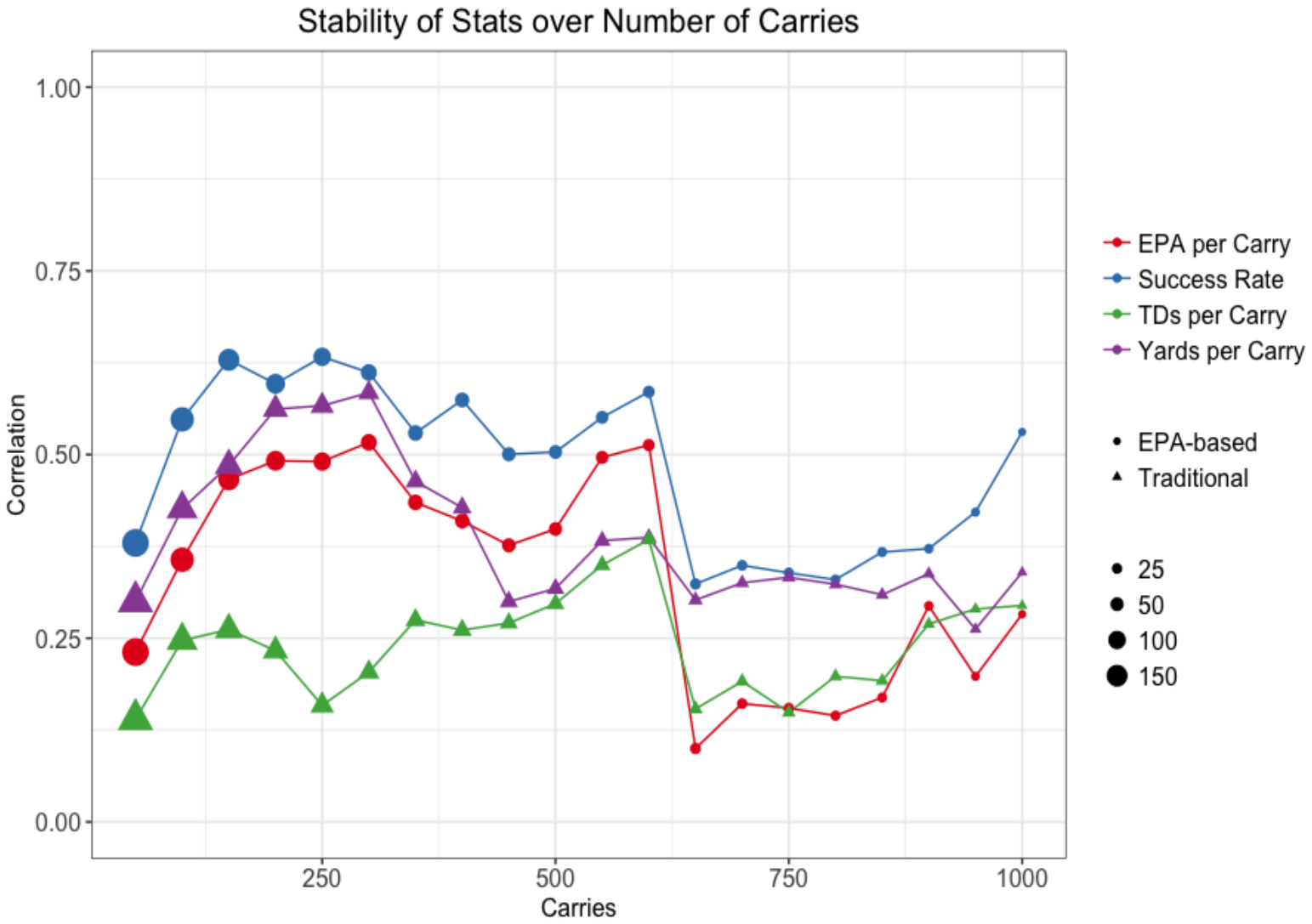




Passing Stabilization



Rushing Stabilization



Consistency Across Seasons

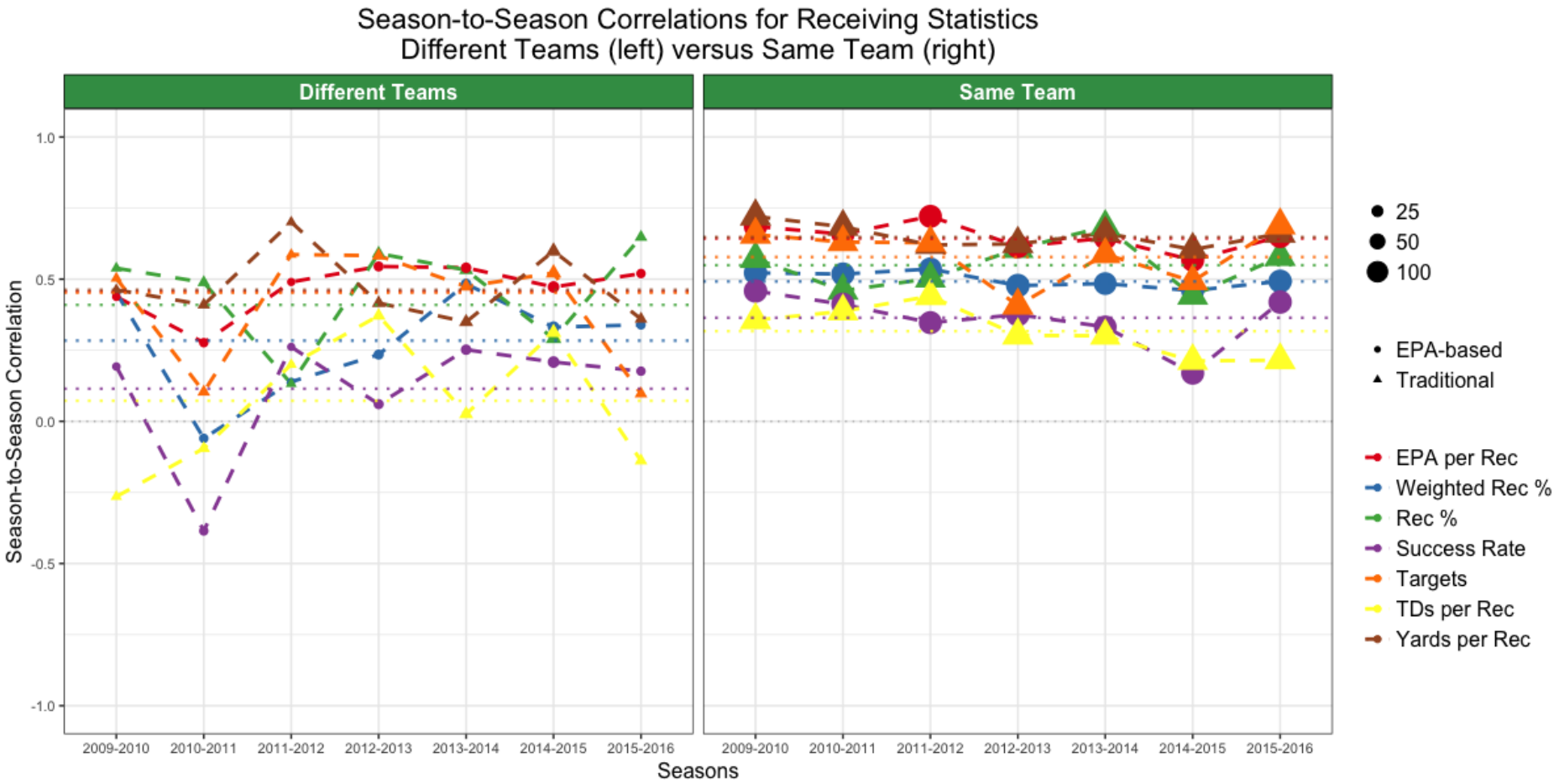
Accurate metrics of player ability **should be consistent** over time

Check the correlations between player seasons in two ways:

- Same team in both seasons
- Different team in each season

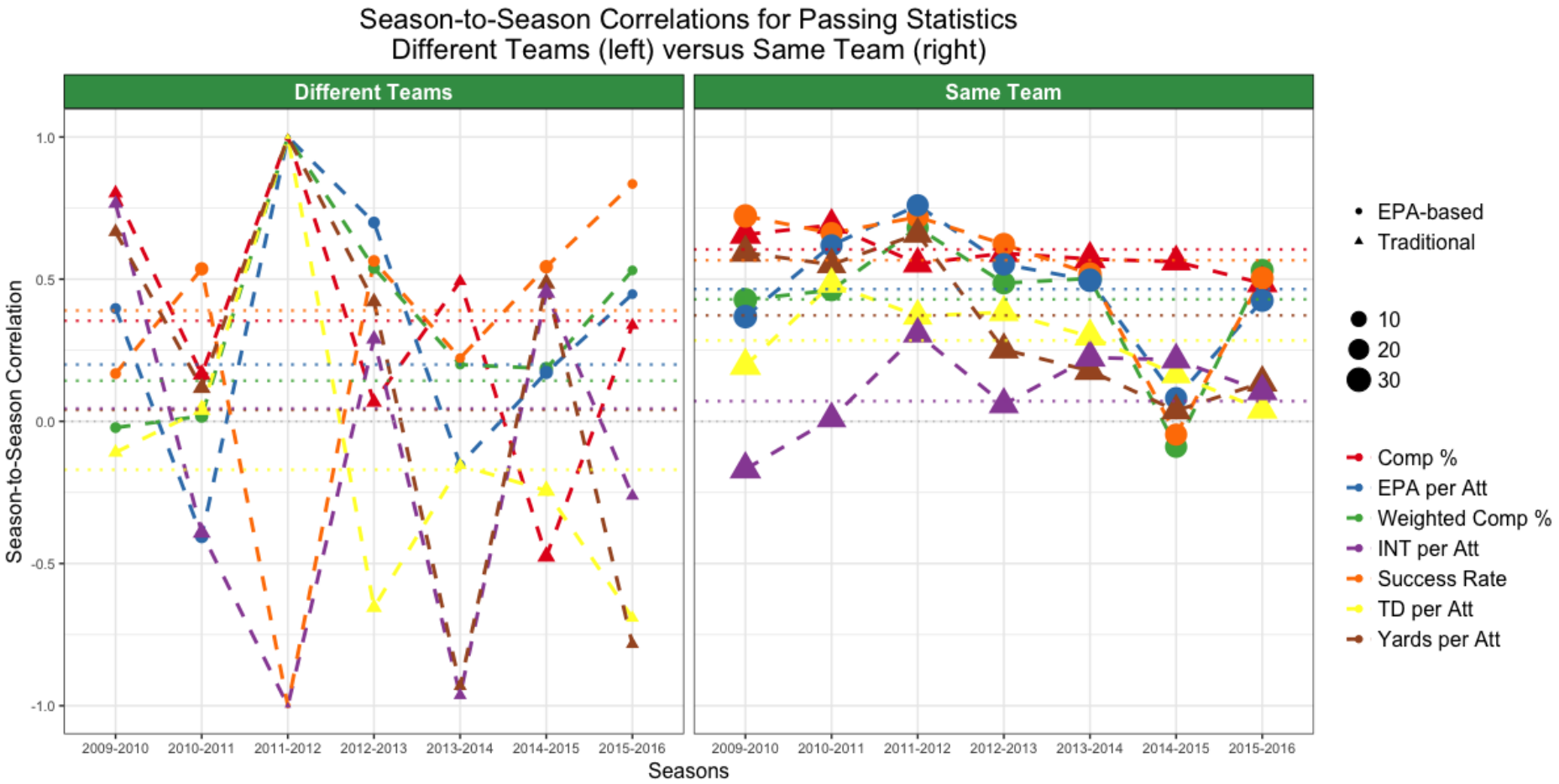
Ideally, measure of player's ability should be independent of team

Receiving Correlations



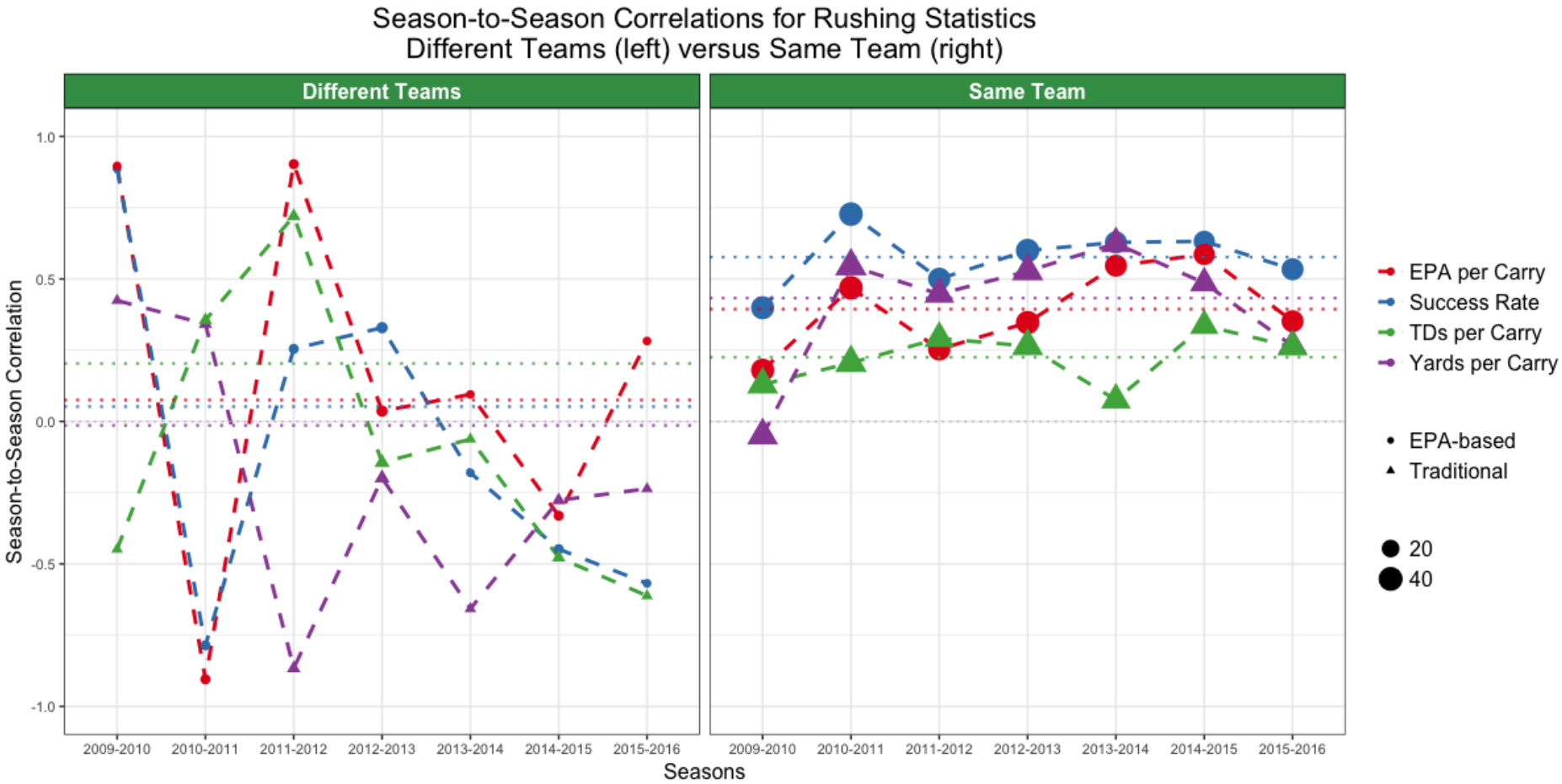
High level of consistency for receiving stats

Passing Correlations



Drop between 2014-2015 led by Peyton Manning and Tony Romo

Rushing Correlations



Clearly **team dependent** (O-line?), e.g. DeMarco Murray:

DAL in 2014 - Success Rate = 43.16%

PHI in 2015 - Success Rate = 34.38%

Recap

- Traditional metrics do not always properly evaluate a play
- Using `nflscrapR` we can calculate EPA based metrics (as well as WPA) and view player production across the field
- Passing is more efficient than rushing from an EPA view

Recap

- Traditional metrics do not always properly evaluate a play
- Using `nflscrapR` we can calculate EPA based metrics (as well as WPA) and view player production across the field
- Passing is more efficient than rushing from an EPA view
- Receiving:
  - EPA per Rec and Yards per Rec are highly consistent
- Passing:
  - Success Rate and Comp % are more appropriate than yards
- Rushing:
  - Success Rate is most consistent
  - Most difficult to evaluate

Future Work: Isolate Player Contribution

Obviously one player is not solely responsible for EPA

Need to account for the situation (down, yards to go, etc.) and also the teams and players involved

Future Work: Isolate Player Contribution

Obviously one player is not solely responsible for EPA

Need to account for the situation (down, yards to go, etc.) and also the teams and players involved

Mixed model approach (Judge et al., 2015):

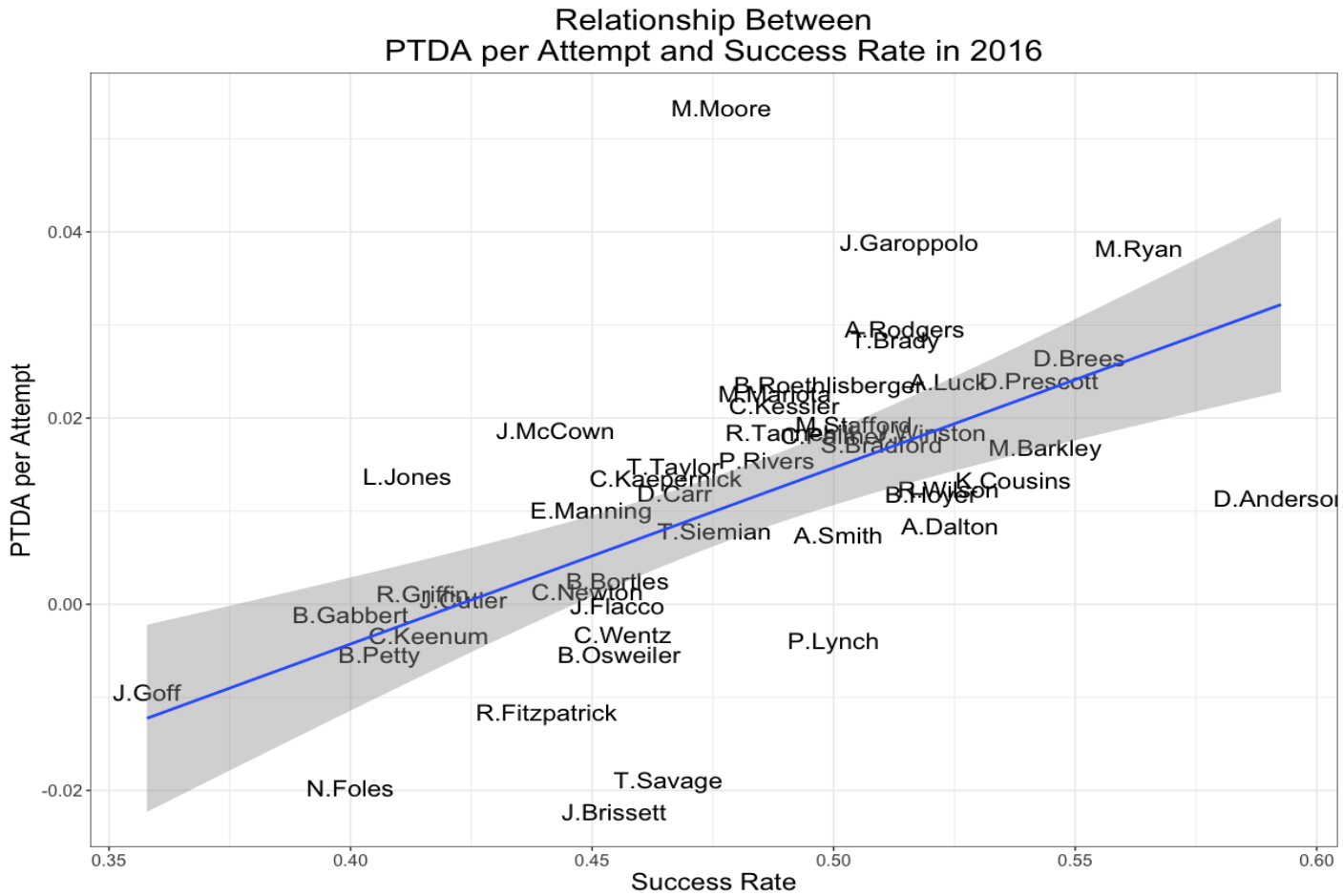
- **Fixed effects** for the situation
- **Random effects** for the individual players, teams

This work will be presented at NESSIS 2017

Future Work: Probability Measures

Multinomial logistic regression model generates probabilities for each of the 7 next score events - can create **new statistics for events**

e.g. **Probability of Touchdown Added (PTDA)**



Tartan Sports Analytics

Website application:

- Developing reports and shiny apps to host on Tartan Sports Analytics Club website
- `https://tartansportsanalytics.com/`

`nflscrapR` **development on Github:**
`https://github.com/maksimhorowitz/nflscrapR`

Follow us on twitter:
Tartan Sports Analytics - @CMUAnalytics
Ron - @Stat_Ron
Sam - @stat_sam
Max - @bklynmaks

References I

[Hastie et al., 2009]. [Carroll et al., 1998].
[Pasteur and David, 2017]. [Goldner, 2017]. [Berri and Burke, 2017].
[Carter and Machol, 1971]. [Burke,]. [out,]. [num,].


Football outsiders.


numberfire.


Berri, D. and Burke, B. (2017).
Measuring productivity of nfl players.
In Quinn, K., editor, *The Economics of the National Football League*, pages 137–158. Springer, New York, New York.


Burke, B.
Advanced football analytics.

References II


Carroll, B., Palmer, P., Thorn, J., and Pietrusza, D. (1998).
The Hidden Game of Footbal: The Next Edition.
Total Sports, Inc., New York, New York.


Carter, V. and Machol, R. (1971).
Operations research on football.
Operations Research, 19(2):541–544.


Goldner, K. (2017).
Situational success: Evaluating decision-making in football.
In Albert, J., Glickman, M. E., Swartz, T. B., and Koning, R. H., editors, *Handbook of Statistical Methods and Analyses in Sports*, pages 183–198. CRC Press, Boca Raton, Florida.

References III


Hastie, T., Tibshirani, R., and Friedman, J. (2009).
The Elements of Statistical Learning: Data Mining, Inference, and Prediction.
Springer, New York, New York.


Pasteur, R. D. and David, J. A. (2017).
Evaluation of quarterbacks and kickers.
In Albert, J., Glickman, M. E., Swartz, T. B., and Koning, R. H., editors, *Handbook of Statistical Methods and Analyses in Sports*, pages 165–182. CRC Press, Boca Raton, Florida.

Expected Points Model

Combined two ideas for weighting plays:

- Score differential - more weight for close score games
- Distance in drives away from next score