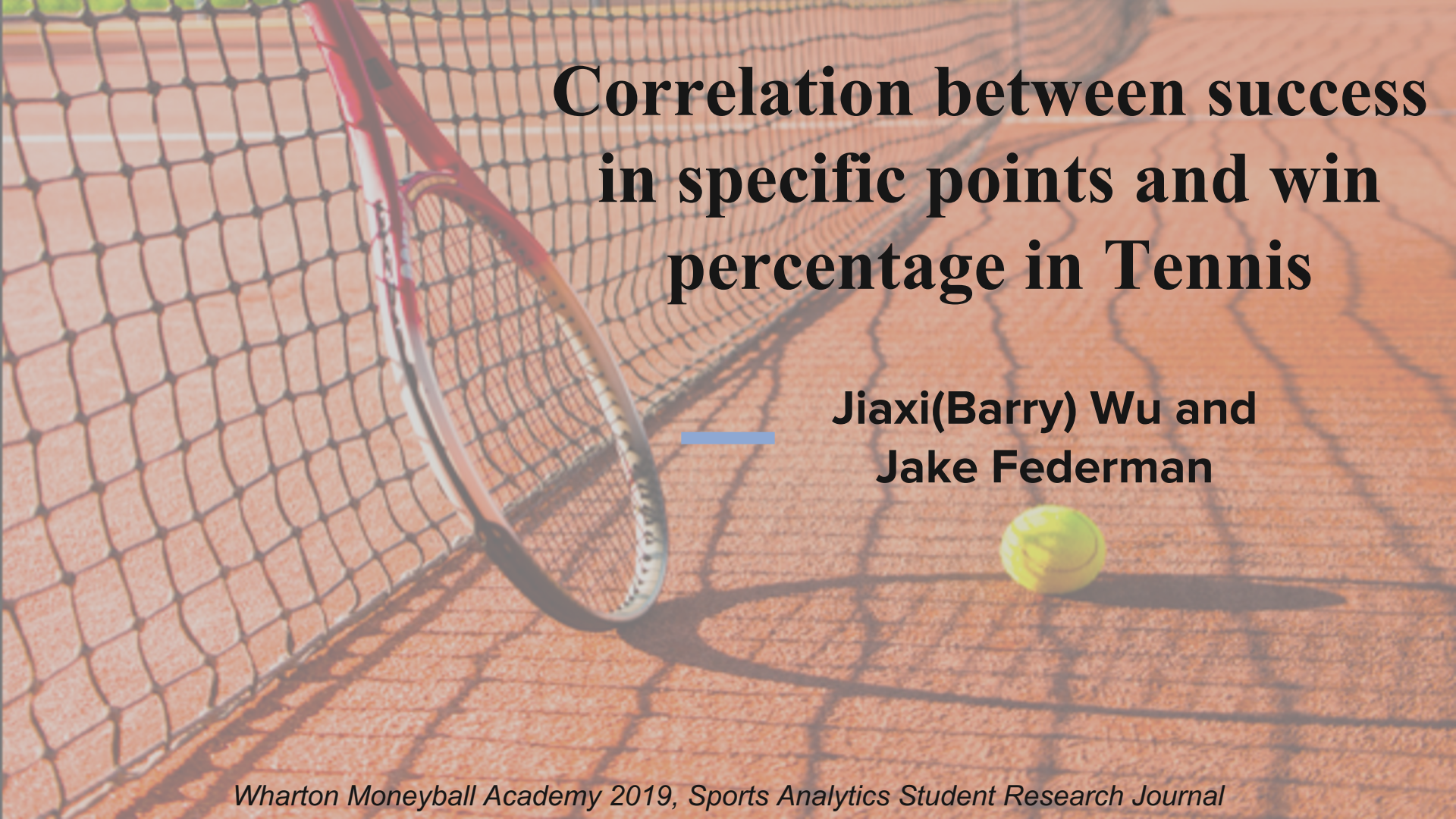


Correlation between success in specific points and win percentage in Tennis

Jiaxi(Barry) Wu and
Jake Federman

Wharton Moneyball Academy 2019, Sports Analytics Student Research Journal

Agenda

1. Discuss what stats correlate well to win percentage, analyze, and ask why the trends we are seeing are occuring
2. Introduce a new metric called the clutch rating
3. Explain a model we built that factors in multiple stats including our clutch rating to determine win percentage
4. Draw conclusions and speculate potential next stats

Data

- Career Grand Slam performance from 2002 to 2019
- Source: www.ultimatetennisstatistics.com

Which stats correlate the most to winning percentage?

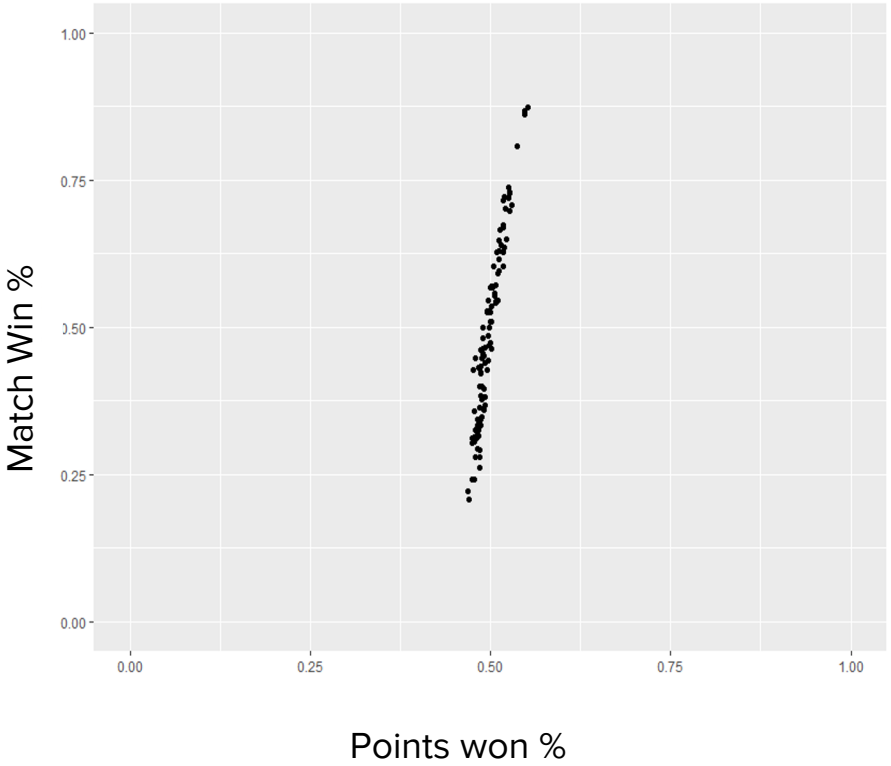
In other words, which stats can explain why some players have been more successful than others?

POINTS WON PERCENTAGE

- Obviously, points won percentage correlates very well to match winning percentage (95%).
  - The player who wins more points should and usually does win the match.

Data from ultimatetennisstatistics.com

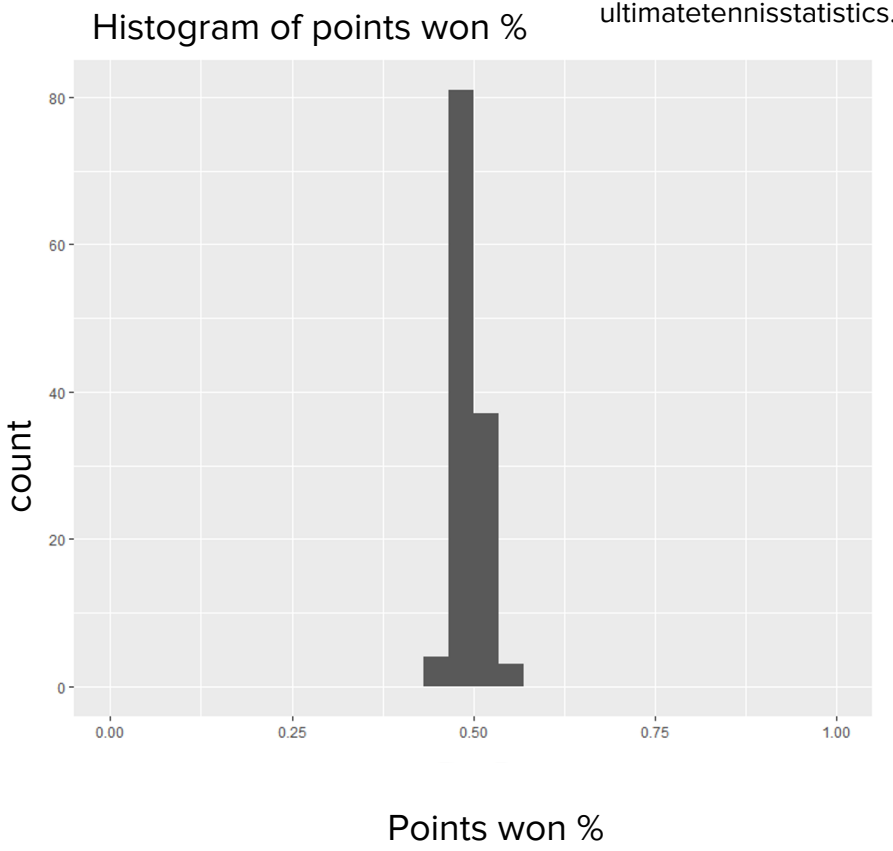
Points won percentage vs match win percentage



Analyzing points won percentage

Data from ultimatetennisstatistics.com

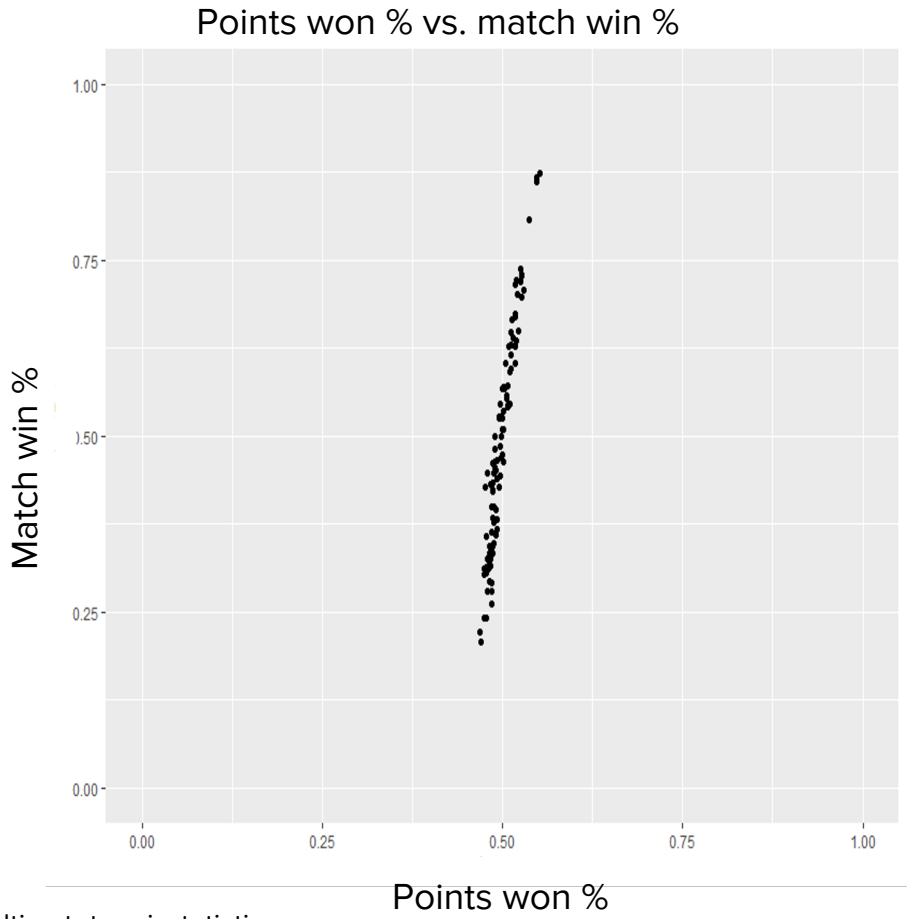
- There is extremely little variation in players' points won percentages from the top players to the bottom
  - Even the best players only win at most 56% of their points.



Analyzing points won percentage and match won percentage

- Yet match win percentage has a very wide range along the points won percentage.
- But how can we understand why that distribution has occurred? How can match won percentage be so spread out when all the players win around the same amount of points?

- Are some points more important than others?



Data from ultimatetennisstatistics.com

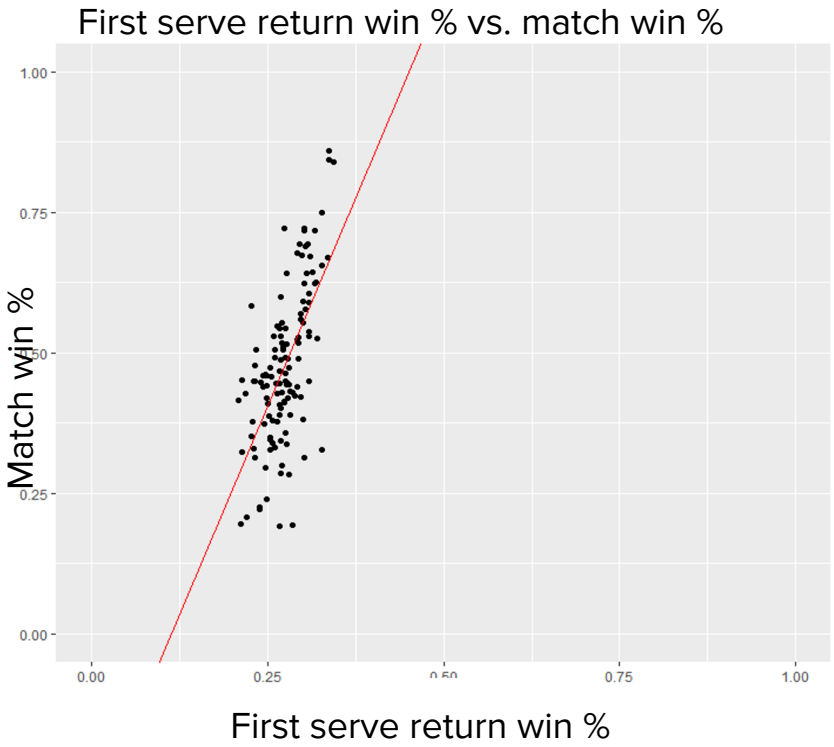
Are some points more important than others?

- Initial hypothesis of some scenarios/stats that could be more important than others
  - First serve win percentage
  - Second serve win percentage
  - First serve return win percentage
  - Second serve return win percentage

Are some points more important than others?

YES.

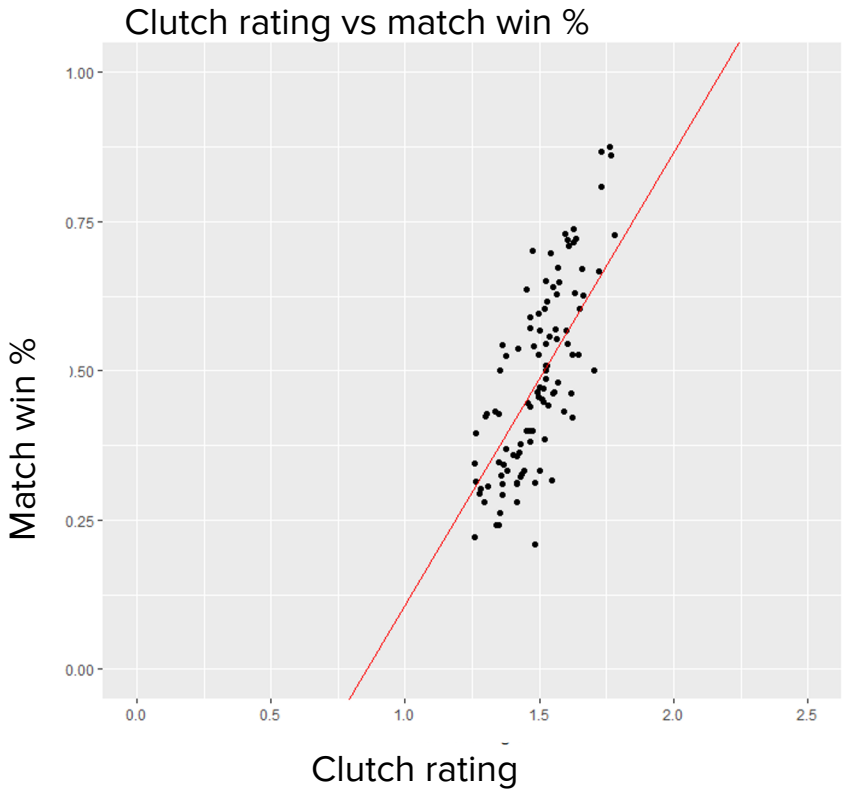
- First serve return win percentage: 0.637
- Second serve win percentage: 0.53
- First serve win percentage: 0.52
- Second serve return win percentage: 0.43



Data from ultimatetennisstatistics.com

Clutch Rating Breakdown

- Clutch Rating = Break point saved percentage + break point conversion percentage
- Correlation: ≈ 0.6



Data from ultimatetennisstatistics.com

Building a Model

- Using the two stats that had relatively strong correlations (first serve return win percentage and clutch rating), we built a model to predict a player's win percentage.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon$$

number of predictors

Testing the Model

The formula:

PredictionStupid=-1.36+1.55*Clutchrating+2.52*First_Serve_Return_Win

(coefficients found via regression)

Please note PredictionStupid is just the name assigned. The prediction is not actually stupid.

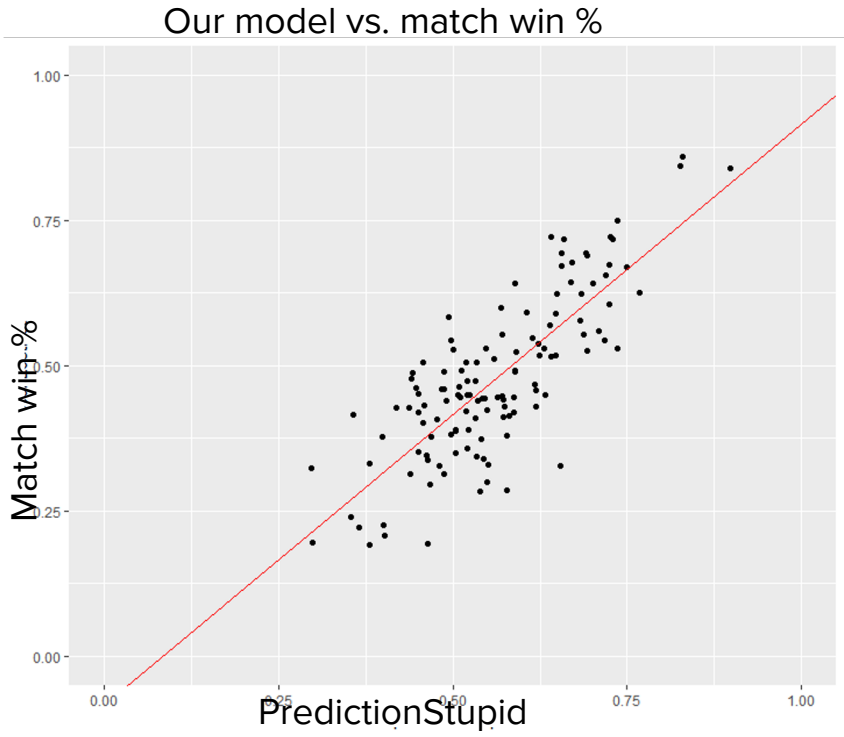
Correlation (r) :0.800

R-squared: 0.6404

RMSE:0.118

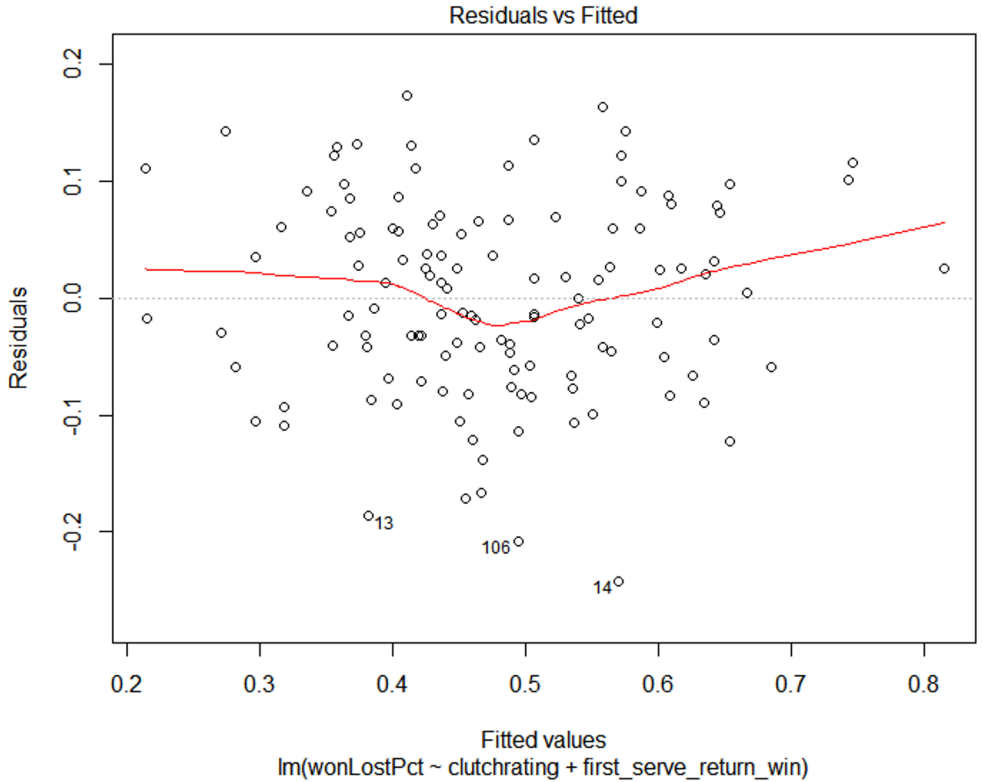
P-value<2.2e-16

Data from ultimatetennisstatistics.com



Residual plots

The incredibly random and even distribution of the residuals (left plot) show the accuracy of the



Conclusions

- There is very little separation in the skill levels of professional tennis players.
  - Even the top players only win 56% of their points at most
- Despite this, the top players are able to win almost 90% of their matches because of their great performance in specific situations
  - The most important of these scenarios are break points and first serve return points
- Our model, which factors in both of those scenarios, correlates very well to match winning percentage (0.8)

Potential Next Steps

- Find a way to eliminate or correlate more the outliers
- Make the model predictive
- Work with a larger data set
  - I.e. work with data from all tournaments
  - Players from all eras
    - And that would require some adjusting for era, as the game has changed over time

THANK YOU!!!

Any questions?