Measuring Goaltending Performance using Bayesian Analysis and Expected Goals

Cole Anderson

@crowdscoutsprts

crowdscoutsports.com

cole92anderson@gmail.com

github.com/C92Anderson/xG-Model

Who am I?

Data Scientist at DigitasLBi

Client work: behavioral & predictive models, ecommerce, technology, retail

Website: crowdscoutsports.com

Previously antitrust economist, consulting during 2012-13 NHL lockout

Former AJHL, BCHL, NCAA* goaltender

*D3

DigitasLBi

What am I talking about?

Create a metric to isolate goaltender performance

Can we predict future performance?

Adjusting for shot quality and rebounds helps *some*

Bayesian framework helps *more*

I. Developing & Applying an Expected Goals Model

What is the probability that this shot will result in a goal?

What responsibility is the goaltenders?

Nature of the Data

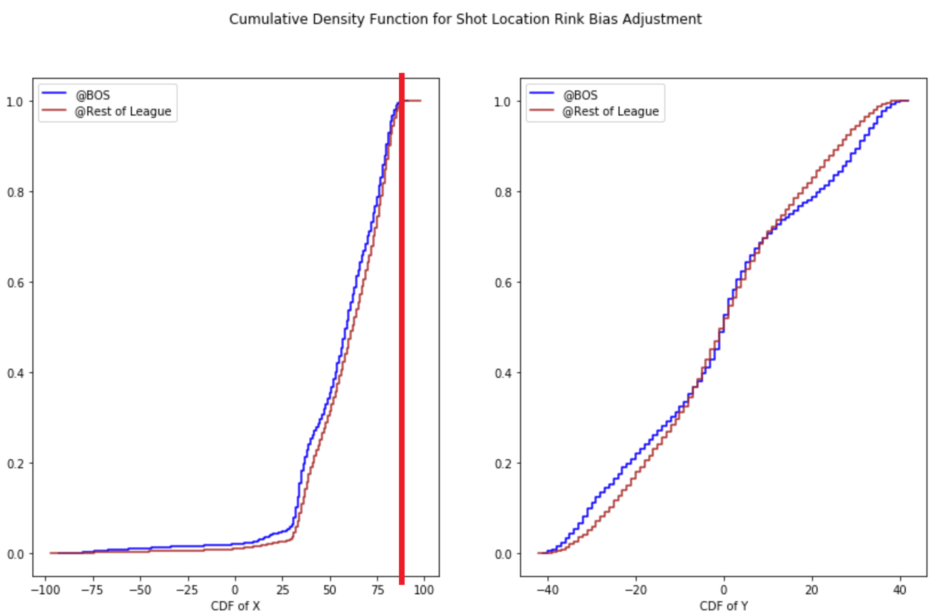
What is the probability that this shot will result in a goal?

- Play-by-play data only captures ‘outcome’ events
- These events are result of complex and fluid game
- “watching a passing train by looking at sparks off the rails”
- Deep learning models will benefit from expanded ’input data’
- In meantime with ‘outcome data’ logistic regression okay

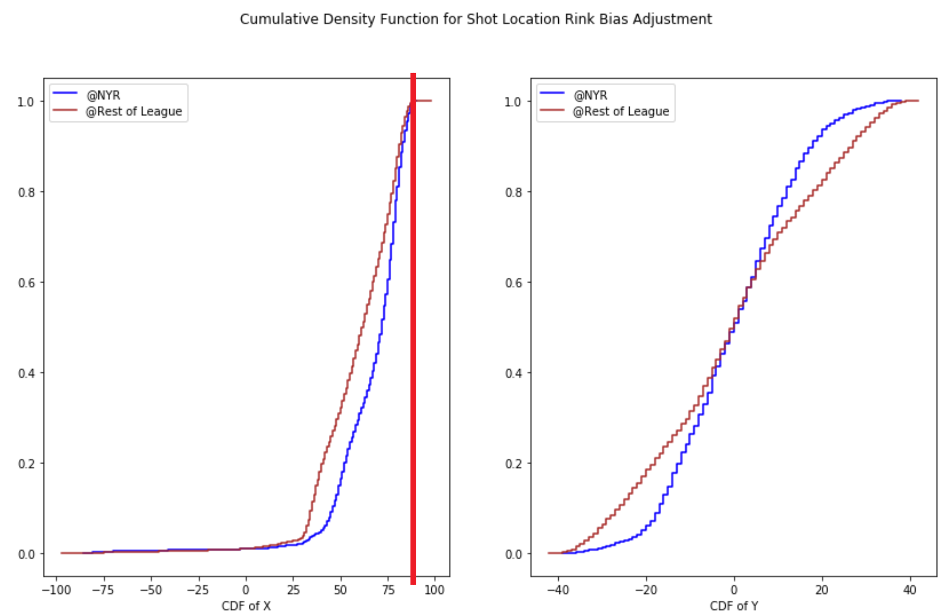
Scorer Bias Ranges from Rink & Season

Coordinates of shots are responsibility of home rink scorer – these can be biased

BOS, 2009-10 – Home v Away

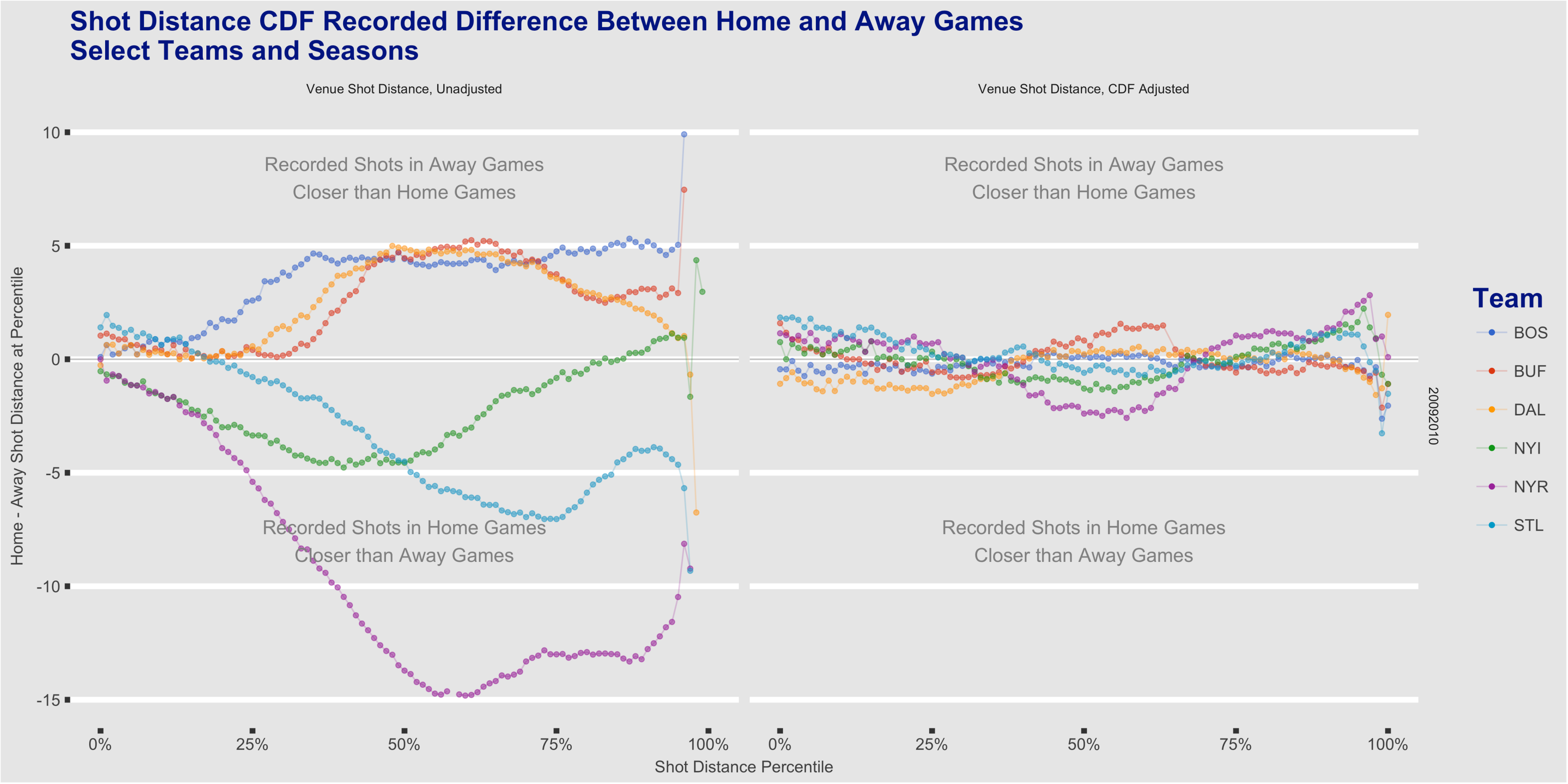


NYR, 2009-10 – Home v Away



https://arxiv.org/pdf/1412.1035.pdf

Scorer Bias Ranges from Rink & Season



Scorer Bias Ranges from Rink & Season

What other team or scorer biases exist that haven’t been discovered and/or can’t be normalized?

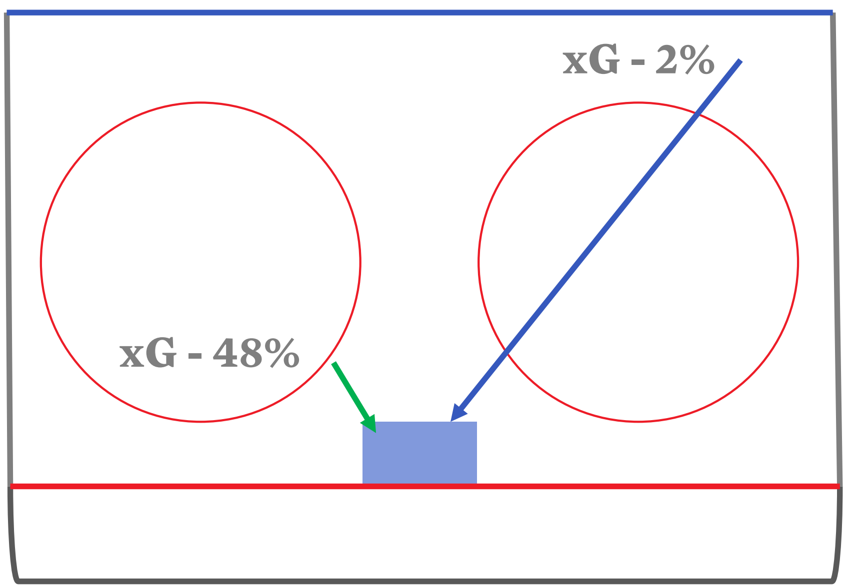
II. Rebound Control

How do we handle the (perceived) influence goalies have on rebounds?

Expected Goals in Action

Rebounds are a problem for analysts and goalies alike

Is this a ½ an xG against?



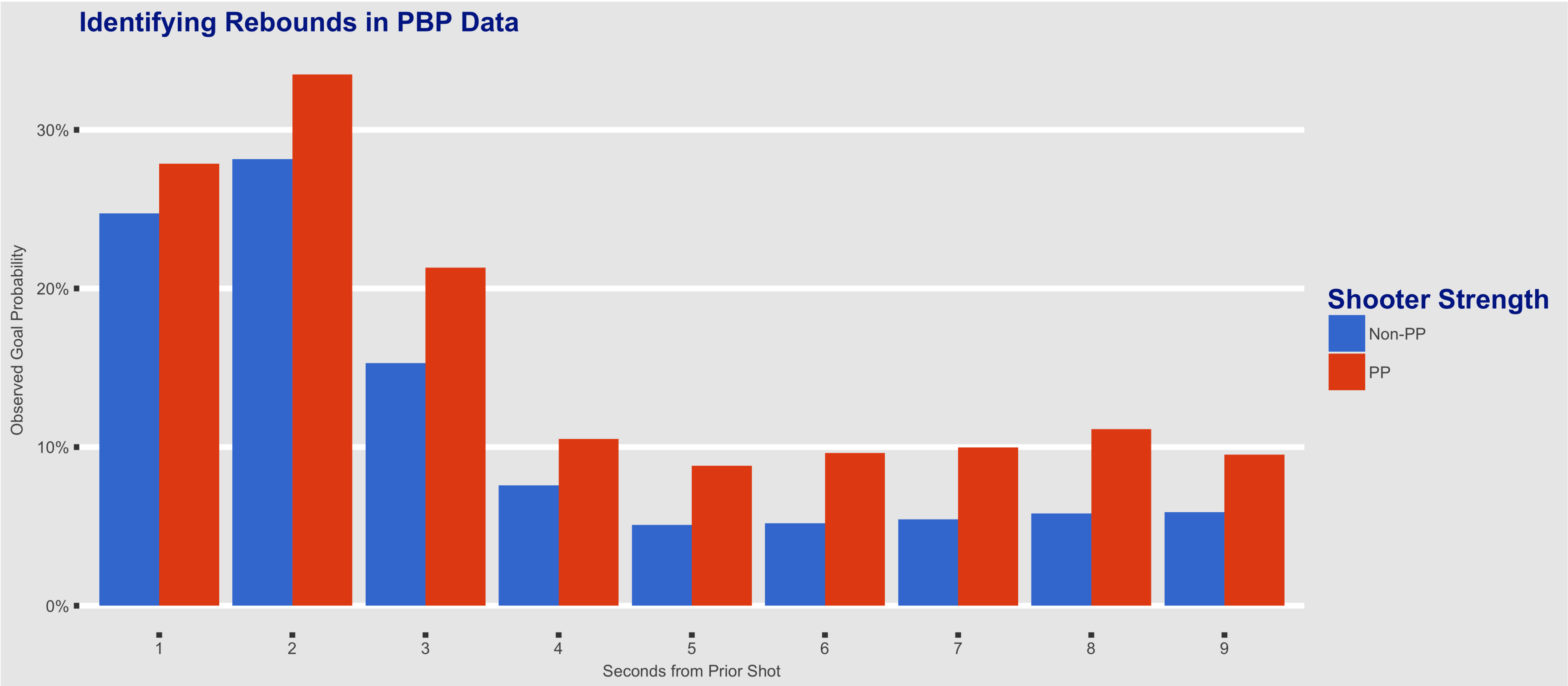
xRebound Model - Overview

Rebounds are a problem for analysts and goalies alike

- Tough to model (not a targeted event)
- Tough to prevent (**** happens, no skill over expected)
- Tough to differentiate (they are all dangerous)
- Tough to summarize (they are multiplicative, not additive)
  - https://medium.com/@dannypage/expected-goals-just-don-t-add-up-they-also-multiply-1dfd9b52c7d0

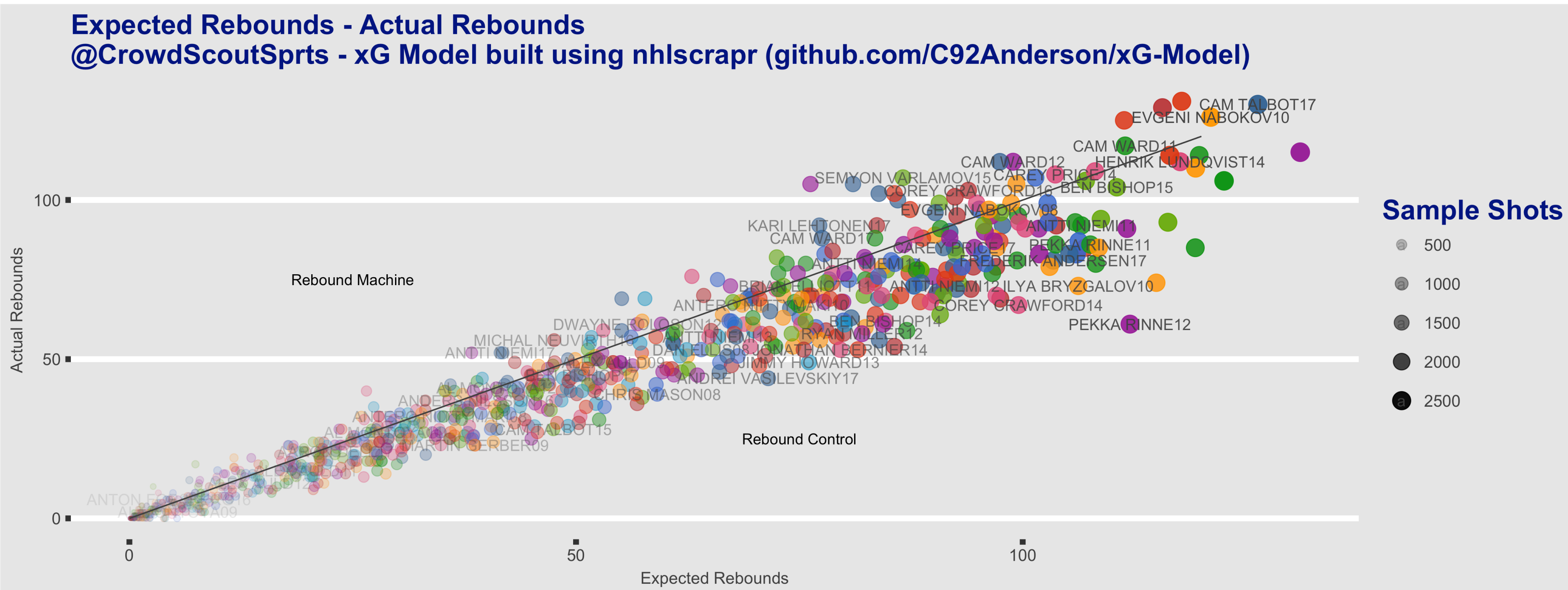
Tough to model (not a targeted event)

What is a Rebound?



Tough to prevent (little observed skill over expected)

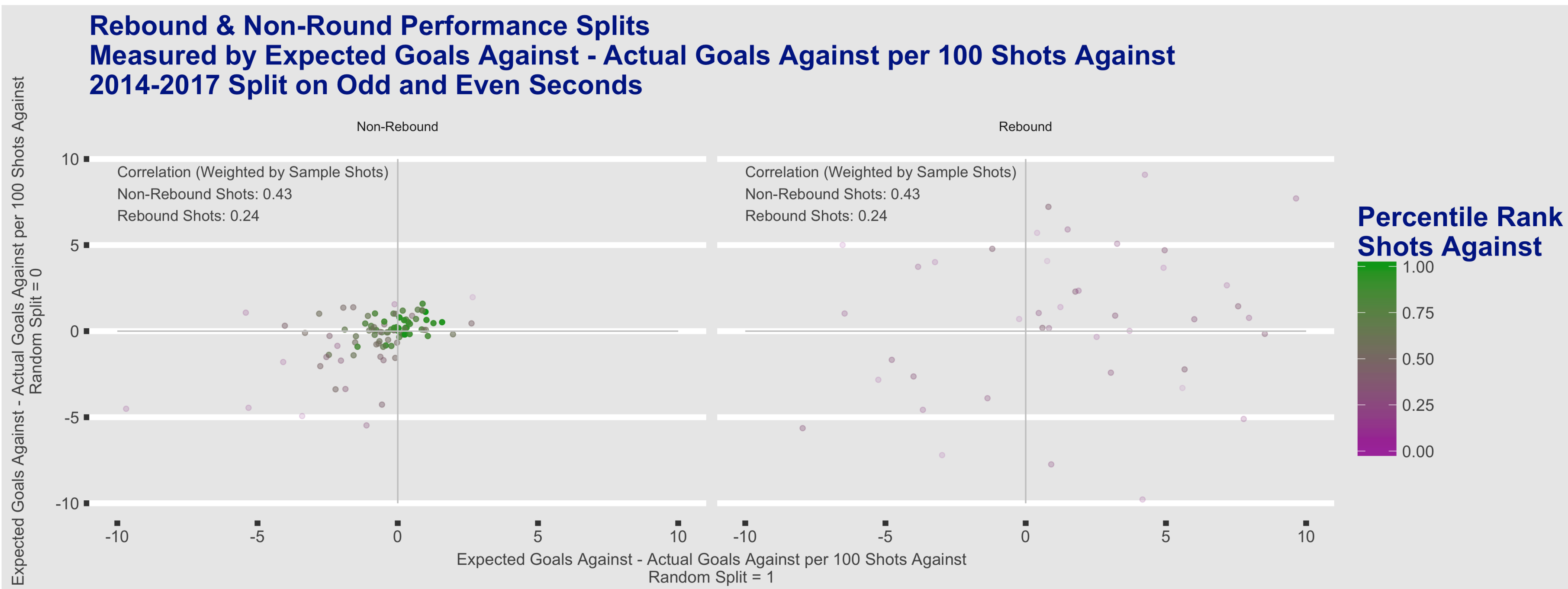
They happen to everyone



Tough to differentiate (less skill stopping them)

Non-Rebound Shot xGA – GA / 100 Shots Intra-season Correlation: 0.43

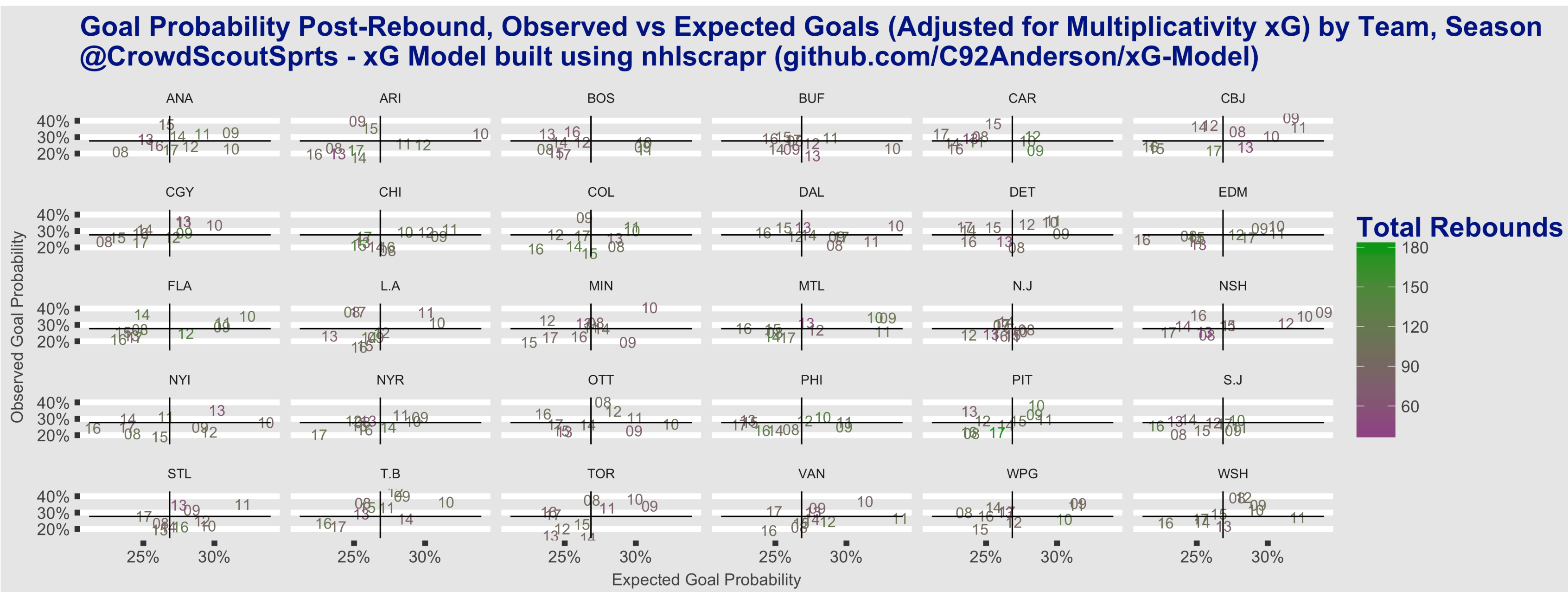
Rebound Shot xGA – GA / 100 Shots Intra-season Correlation: 0.24



Tough to summarize (multiplicative, not additive)

Rebounds end up in a goal (in same flurry of shots) **27.6%** of the time

Adjusting Rebound xG for multiplicatively infers goal **26.9%** of the time



Rebound Control

- Determine what a rebound is, not a ‘clean’ target event
- Consistently preventing them (relative to peers) is tough
- Consistently stopping them (relative to peers) is tough

Test: Simulate rebounds, treat them as ¼ of a goal and test relative predictivity

III. Bayesian Application

What is the base rate?

What is the strength of evidence?

How relevant is this evidence?

Empirical Bayes Estimation

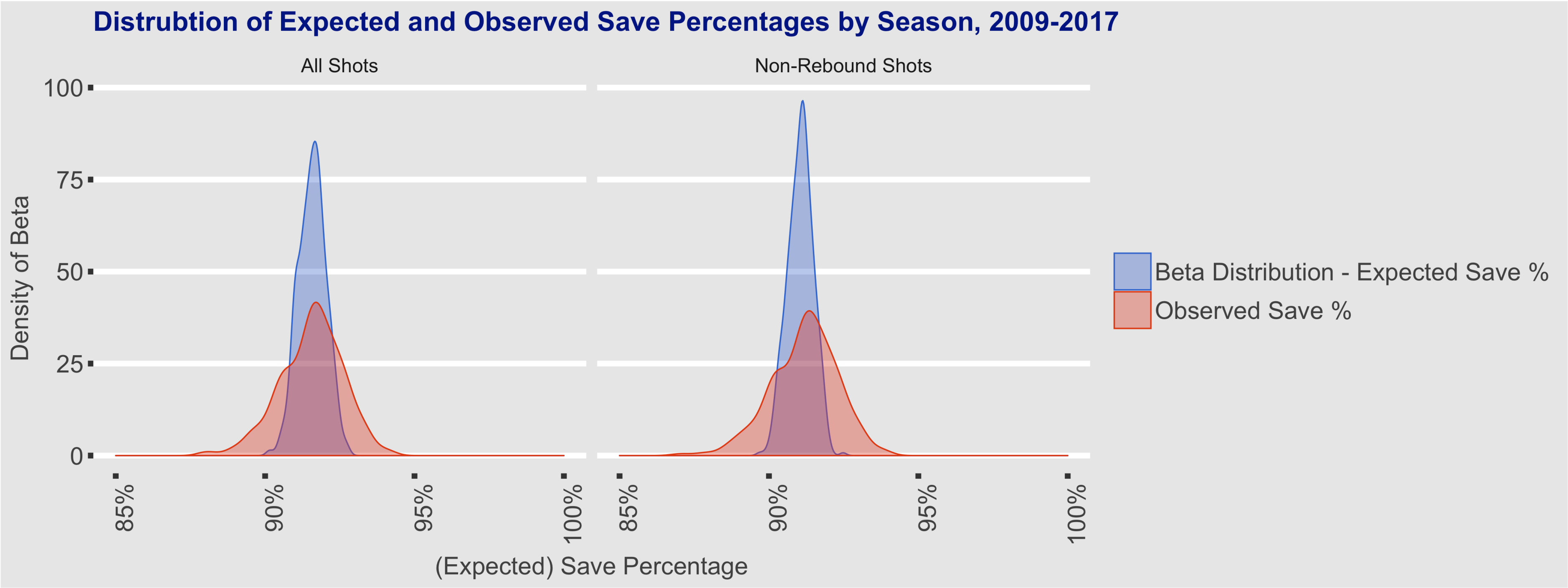
$$\frac{9}{10} \neq \frac{90}{100}$$

Filtering can be arbitrary & removes information

We tread carefully deciding how much prior information to include (if any)

Beta Distribution

Create prior distribution of Expected Save %



The Matt Murray Question

Is Matt Murray:

a) an unbeatable cyborg?

b) a top 10 NHL goalie?

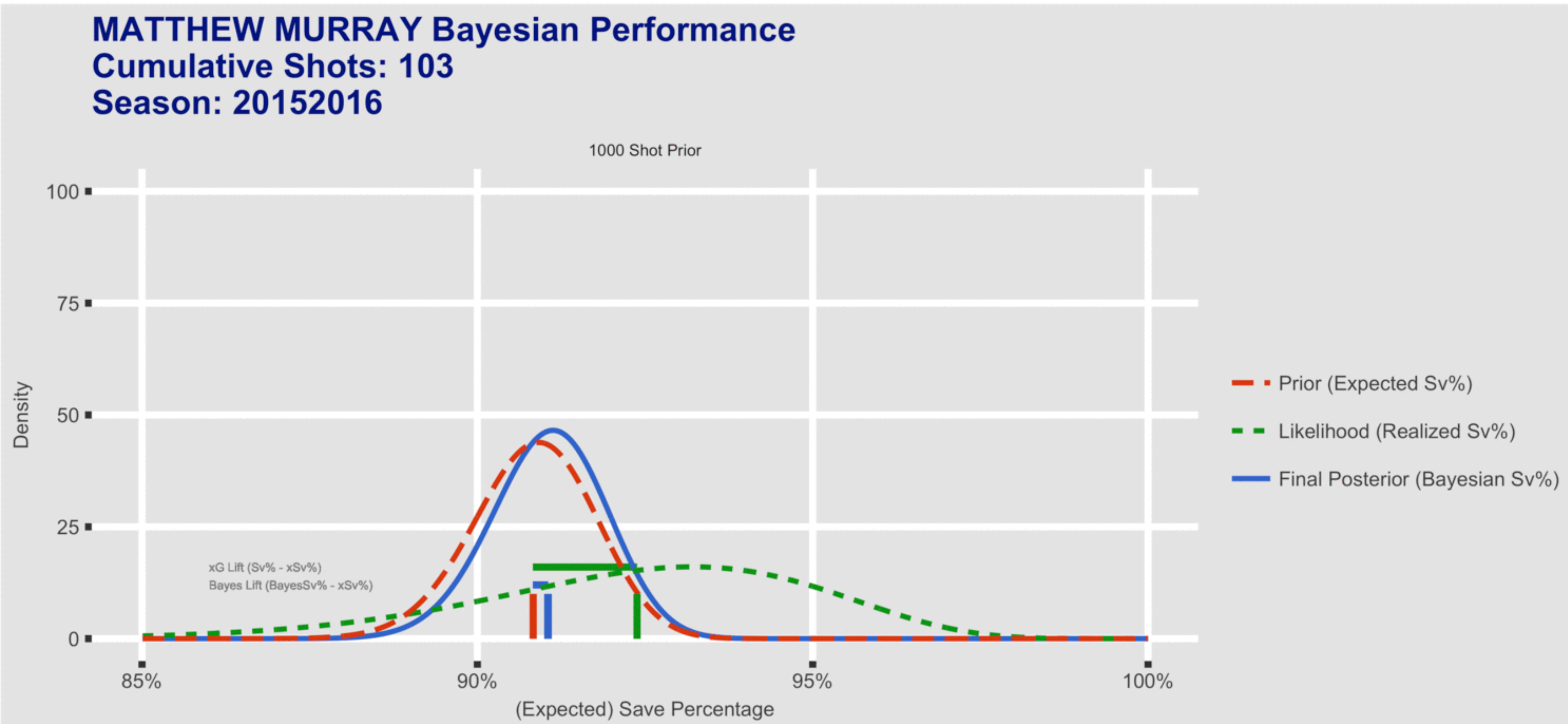
c) an average and fortunate NHL goalie?

d) Marc-André Fleury’s former caddy/backup?

e) leave me alone.

The Matt Murray Question

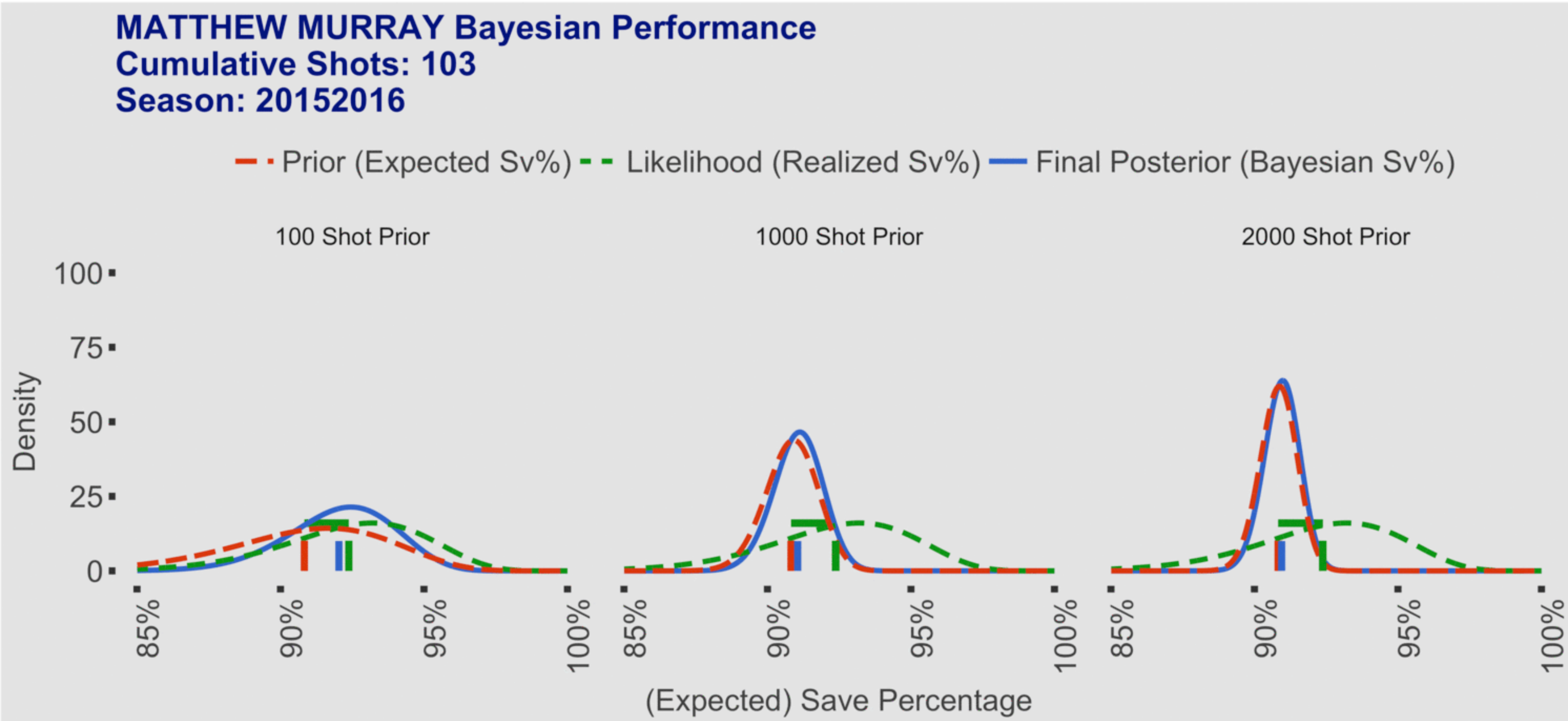
Q: How much has Murray's Sv% exceed his Expected Sv% so far?



www.greaterthanplusminus.com/2013/10/bayesian-approach-to-analyzing-goalies_7222.html

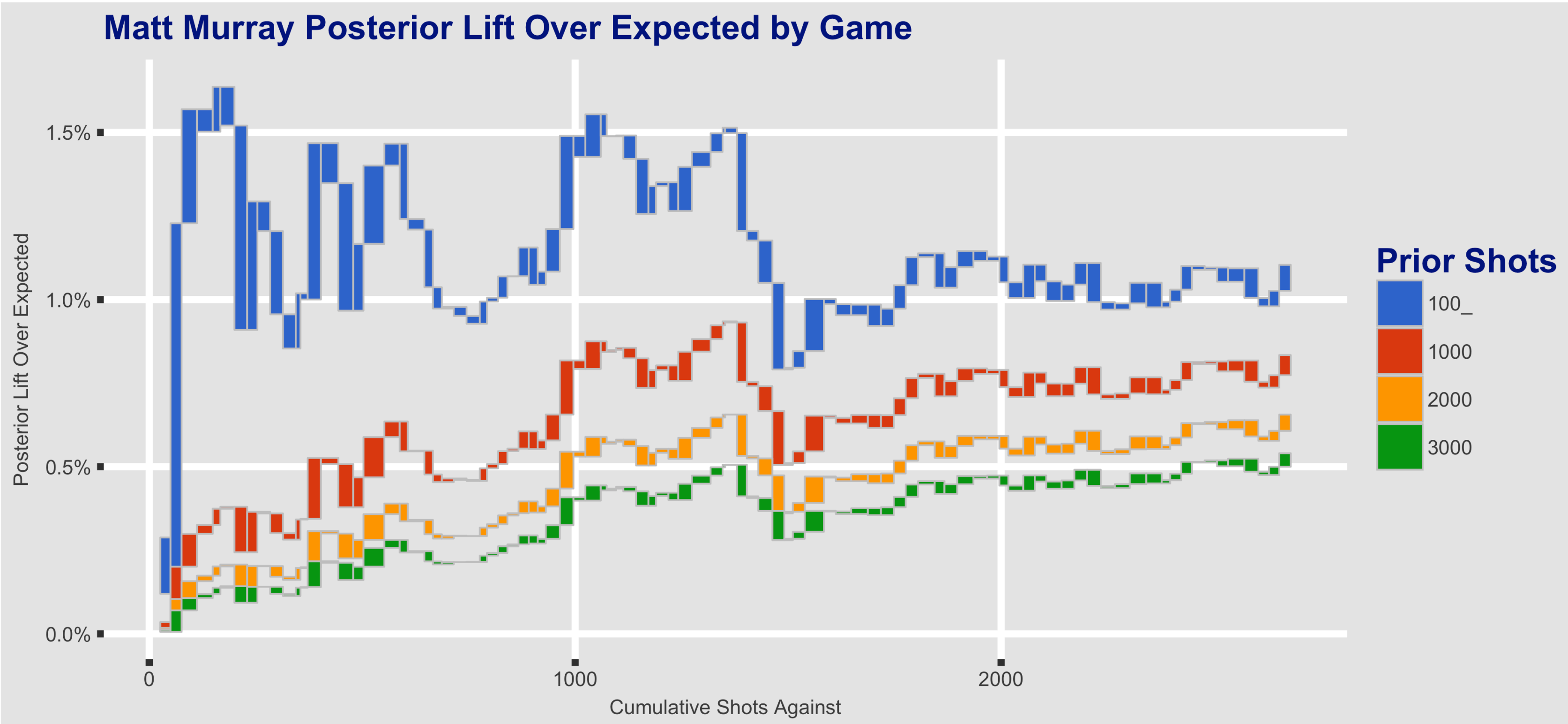
The Matt Murray Question

A: It depends on the strength of the prior



Prior Matters

Signal (Posterior Lift Over Expected Sv%) is influenced by prior



Prior Selection

What prior achieves aim of predicting best?

IV. Lean on the Machine

What metrics project best?

What parameters project best?

The Setup

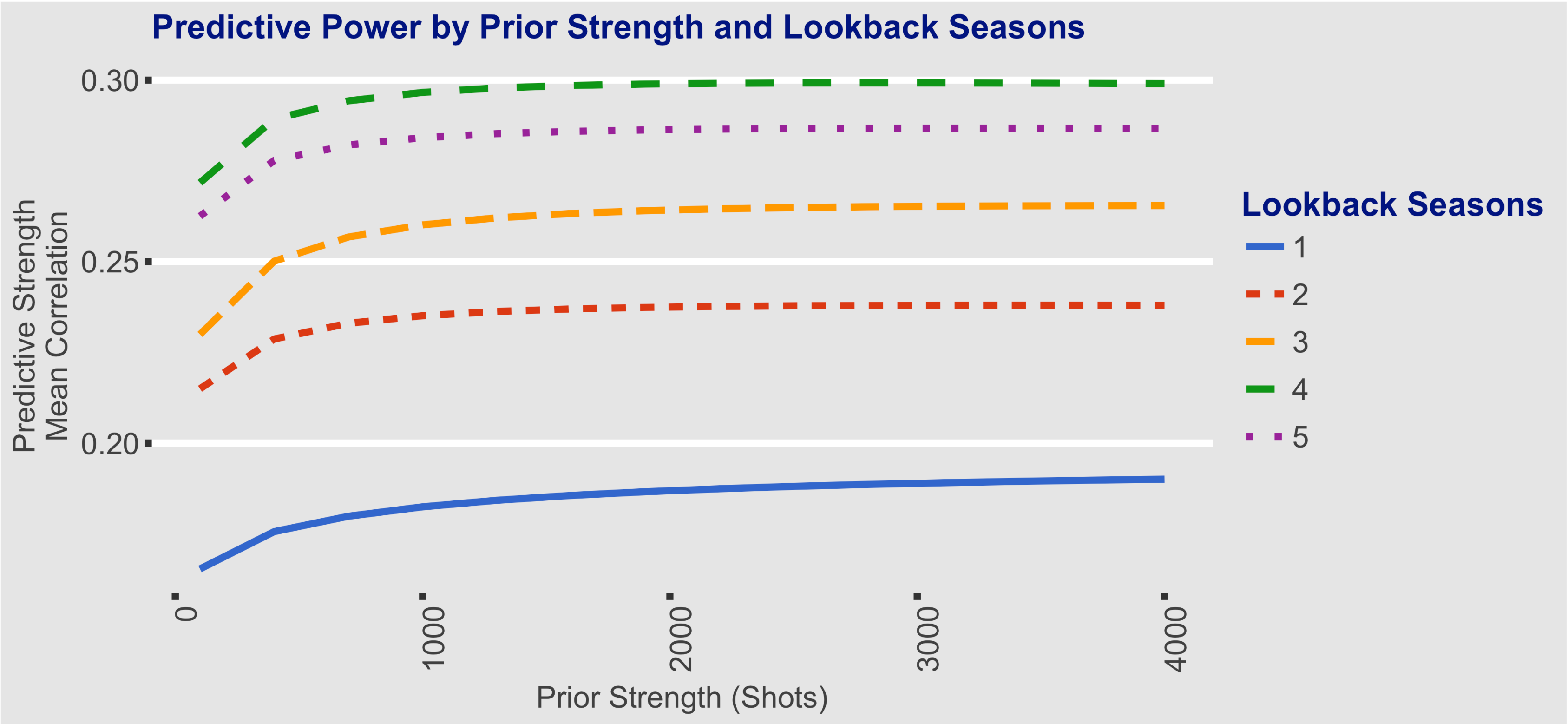
- Test metrics (testing rebound adjustment for each)
  - Raw sv%
  - xG – GA Lift / Shot
  - Bayesian Lift
- Project and compare 6 past seasons, 2011-2017
- Test look back (marcel) years
  - 1 to 5 years
- Test age regression
  - Regress old players more based on 2 parameters
- Test prior strength
  - 100, 400.. 2500, 3000 shots

What combination of metrics and parameters predict best?

2100 Combinations of Parameters

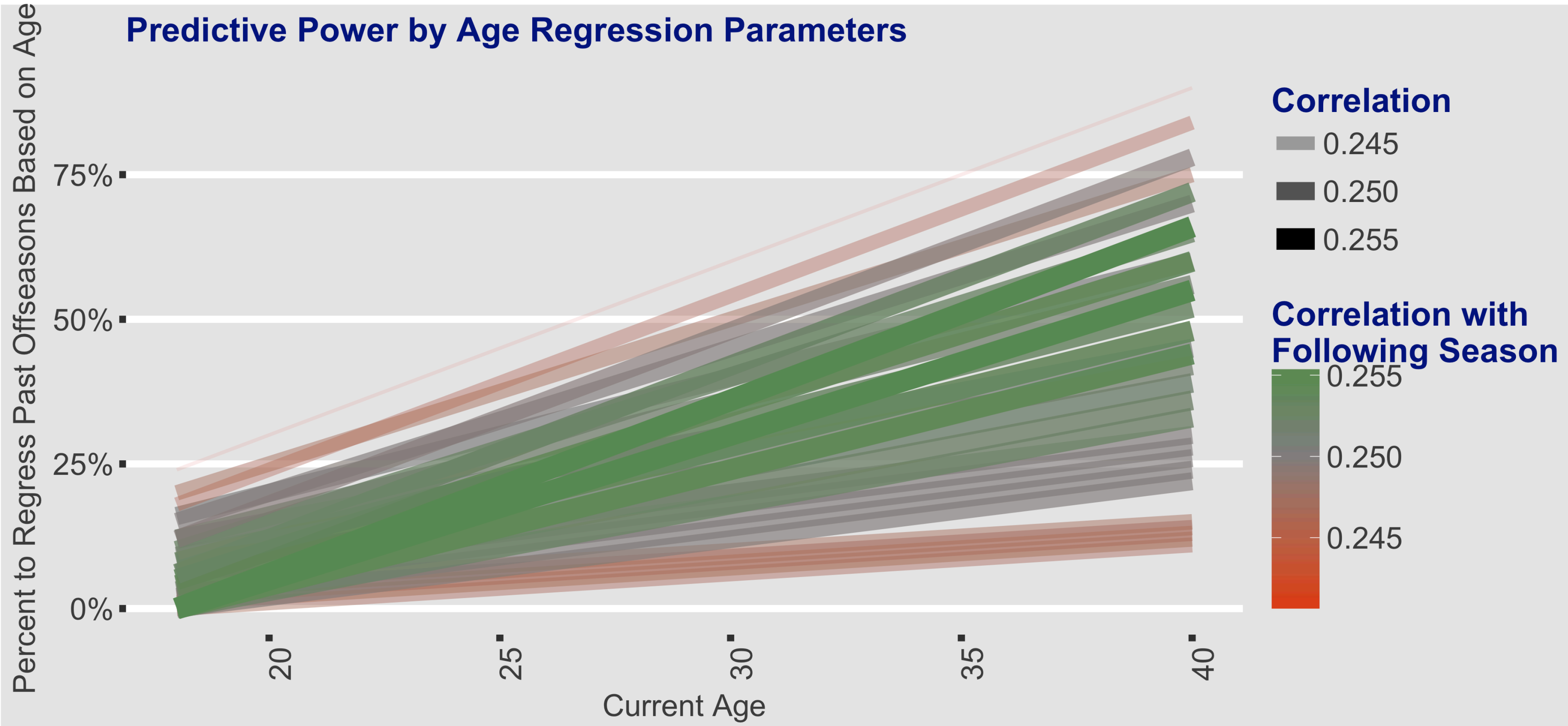
Prior Strength

Predictive power levels of with 4 season lookback & ~1500 shot prior



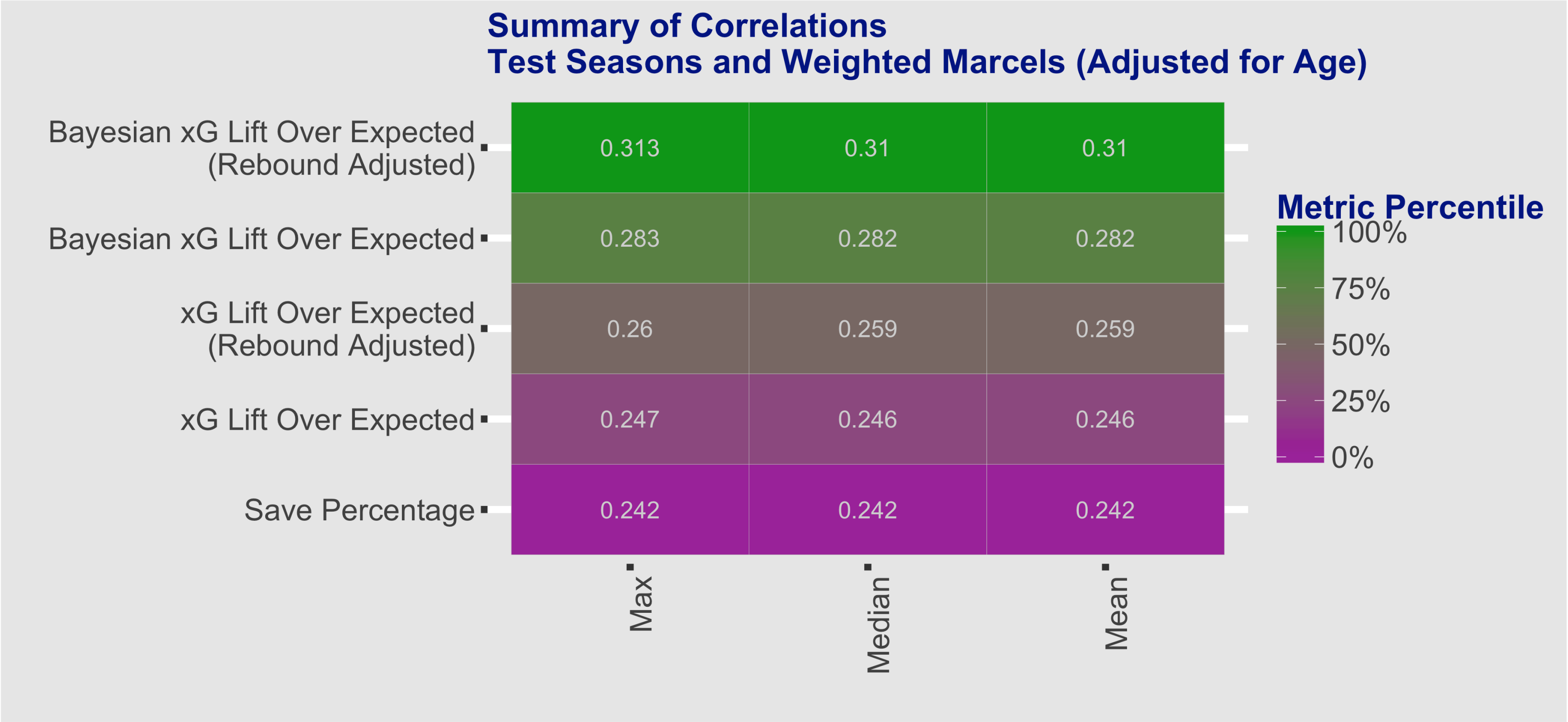
Age Regression

How do we best adjust for aging?



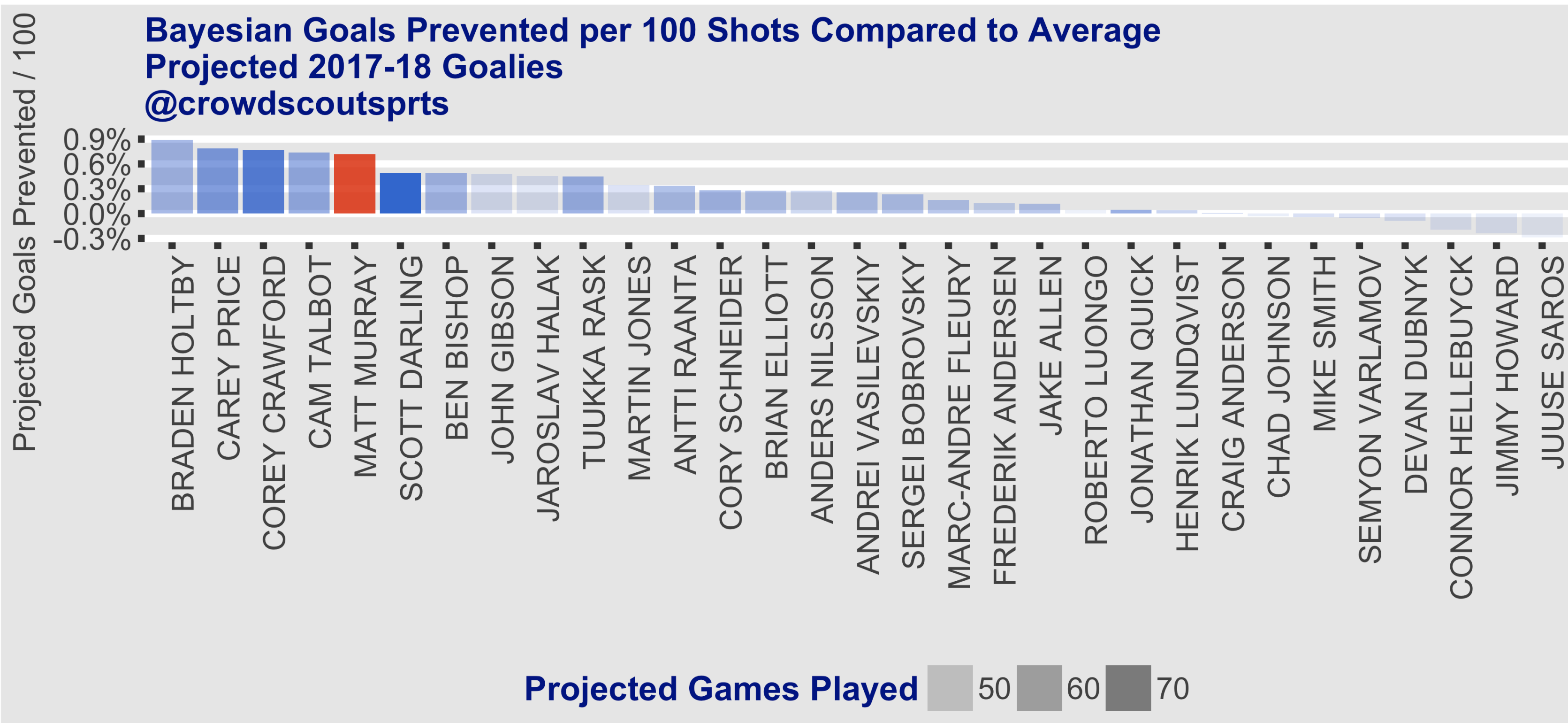
Results Summary

Bayesian xG Lift Over Expected Adjusted for Rebounds performed best



Projecting 2017-2018

Definitive Matt Murray Ranking



Summary

- Bayesian framework increases predictivity
- Modeling rebounds and adjusting for them increases predictivity
- Still going to be wrong a lot
  - xG models are biased, difficult to control for team-level issues
  - Distribution of true talent is tight

Next Steps & Improvements

- Expand sample (i.e. AHL data)
- Updated xG and xRebound Model
  - Always evolving data, feature engineering, and modeling
- Testing different frameworks (i.e. non-linear age)
  - Model follows, doesn't lead aging

Thank you!

Cole Anderson

@crowdscoutsprts

crowdscoutsports.com

cole92anderson@gmail.com

github.com/C92Anderson/xG-Model