

# Value of a(n) (Average) Zone Start

Michael Schuckers

St. Lawrence University  
Statistical Sports Consulting, LLC  
schuckers@stlawu.edu  
@SchuckersM  
@EmpiricalSports

Special thanks to Jim Curro, Dennis Lock, Matt Generous, Jake Hurlbut  
& Chris Wells

September 10, 2016

# Zone Start

Def'n: A Zone Start (ZS) is starting a shift outside of the neutral zone

Goal: Assess average impact of starting a shift in Off Zone (assuming symmetry the negative of impact for starting a shift in the Def Zone)

# Motivation

For 2015-16

David Backes (STL)

25% (312) starts in Off Zone(OZ)

41% (509) starts in Def Zone(DZ)

45 Points, CorsiRel -0.9 %

Bobby Ryan (OTT)

40% (456) starts in OZ

25% (285) starts in DZ

56 Points, CorsiRel -2.1 %

How to compare (and adjust)?

5v5 Data from Corsica.hockey

# Introduction

Vic Ferrari began this area

Fenwick, Purdy (JLikens) continued this work

Hoag, JaredL, DJohnson made attempts

Eric Tulsky: nice summary

<http://nhlnumbers.com/2012/11/5/>

zone-start-adjusted-corsi-corrections-faceoffs

Lots of work

# The Crux

How to isolate effects given:

- Players for

- Players against

- Score effects by period effects

- Rink biases

- Home Ice effects

- Etc.

Even Strength only, more complicated for Special Teams

# Multiple Regression Event Models

## Event level Player evaluation models

- ▶ Macdonald (2012)
- ▶ Gramacy, Taddy, Jensen (2013)
- ▶ Schuckers and Curro (2013)
- ▶ Thomas, Ventura, Jensen, Ma (2013)

Statistics Literature falls under Regression Adjusted Plus-Minus (RAPM)

# Multivariate Regression

$$\begin{aligned}
 Y_i &\sim \mu + \theta ZS_i + \nu Rink_i & (1) \\
 &+ \sum \beta_j X_{Hij} - \sum \beta_j X_{Aij} & (2) \\
 &+ \gamma(HS_i - AS_i)I_{|HS_i - AS_i| > 1} & (3) \\
 &+ \delta(HS_i - AS_i)(1200 - t_i)I_{|HS_i - AS_i| > 1 \& Period=3} & (4)
 \end{aligned}$$

Introduction

Statistical Model

Results

Conclusions

For 5v5 event  $i$

$X_{Hij}$  is indicator (0,1) that player  $j$  is on the ice for Home team

$X_{Aij}$  is indicator (0,1) that player  $j$  is on the ice for Home team

$HS_i$  is current home score,  $AS_i$  is current away score

$ZS_i$  is (-1,0,1) if event started in Def, Neu, Off, respectively  $t_i$  is seconds into period of event

# Responses

XG:

+P(goal), - P(goal) for Home, Away respectively

Fenwick:

+1, -1 if Shot, Miss for Home, Away respectively

Corsi:

+1, -1 if Shot, Miss, Blocked for Home, Away respectively

NP20:

P(Home Goal in next 20 sec)-P(Away Goal in next 20 sec)



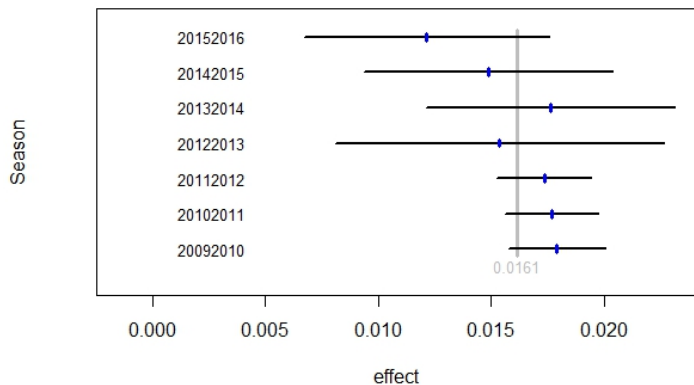
# Some Statistical details

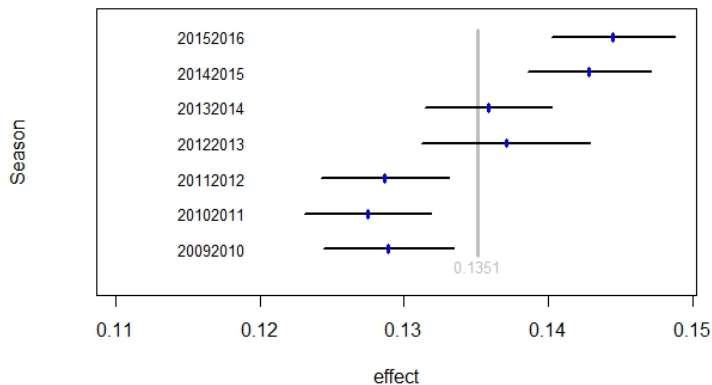
- ▶ Model form = Linear
- ▶ Link function = Identity
- ▶ Estimation via Ridge
- ▶  $\approx 250\text{K}$  event per year
- ▶  $\approx 1100$  players per year

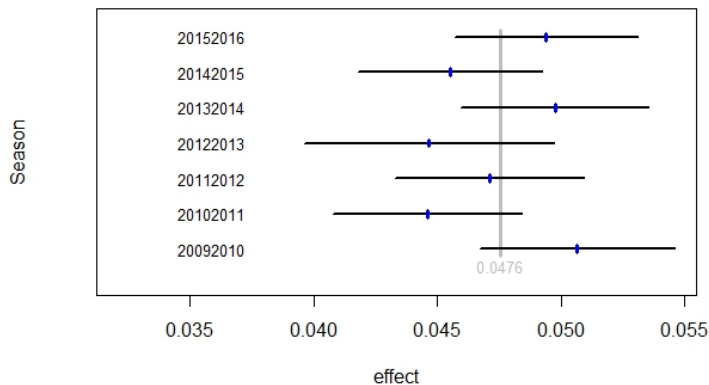
Ridge parameter chosen based upon maximization of out of sample year to year correlation of player effects

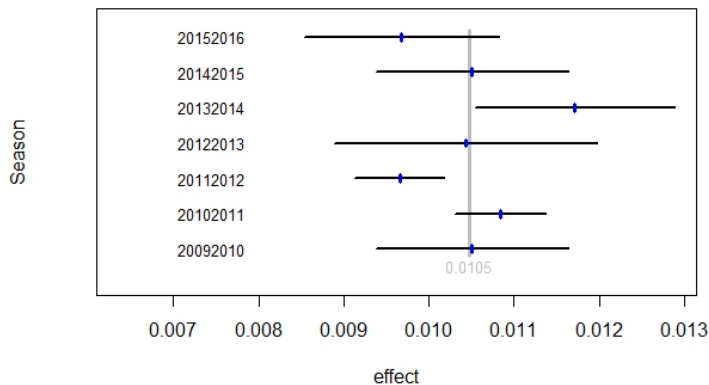
# The Data

- ▶ 7 seasons of data, 2009-10 to 2015-16
- ▶ All 5v5 events (fancier model for special teams)
- ▶ dependent upon NHL RTSS
- ▶ error checking/correction, drop events if an issue

**99% CI for ZS Effect xG in Exp Goal by Shot**

**99% CI for ZS Effect Fenwick In Percent**

**99% CI for ZS Effect Corsi In Percent**

**99% CI for ZS Effect NP20 by Event**

# Estimated Average Impact

## Mean Effects

Response	Effect
xG	0.0161
Fenwick	0.1351
Corsi	0.0476
NP20	0.0105

- ▶ Taking Big Data, event level approach
- ▶ First Order Player Effects, don't have interactions
- ▶ Über WOWY plus
- ▶ Estimates are **averages** per season
- ▶ Get standard errors
- ▶ Making comparable values for player evaluation adj for context



Thank You

`schuckers@stlawu.edu`