Biased Penalty Calls in the NHL

Tim Swartz - Simon Fraser University

Oliver Schulte - Simon Fraser University
David Beaudoin - Université Laval

Ottawa Hockey Analytics at Carleton - January 16, 2016

Introduction

- NHL referee monitoring versus public perception
- related literature
  - Moskowitz & Wertheim (2011)
  - Schuckers & Brozowski (2012)
  - Abreveya & McCulloch (2014)

Data

- 2009/10 - 2013/14 NHL regular seasons: 5,664 matches
- only penalties leading to manpower adv: 42,424 penalties
- scraped game logs from nhl.com

Methodology

- machine learning techniques (e.g. gradient boosting)
- dependent variable $y = 1(0)$ penalty on home (away) team
- logistic regression: $y \sim \mathrm{Bernoulli}(p)$ where $p = \mathrm{Prob}(y = 1)$
- covariates x_1, x_2, x_3, x_4
  - $x_1 \equiv$ total road penalties minus total home penalties
  - $x_2 \equiv$ total road goals minus total home goals
  - $x_3 \equiv$ time in the match when penalty called (0,65)
  - $x_4 \equiv$ team strength parameter (1/0/-1)

Fitted Logistic Model

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.124 + 0.401x_1 - 0.052x_2 - 0.030x_4 - 0.005x_1x_3$$

- next penalty more likely on the road team
- next penalty more likely on the team with fewer penalties
- next penalty more likely on the team having scored more goals
- next penalty more likely on the weaker team
- as matches progress, the penalty differential effect decreases

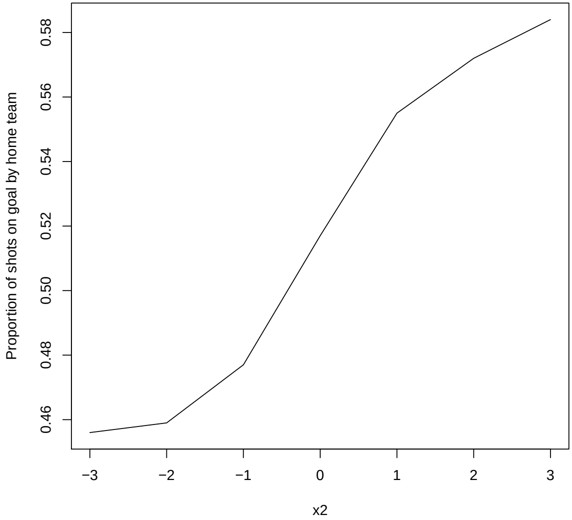
Cause and Effect

Penalty calls are affected by the game situation (x_1, x_2, x_3, x_4) !

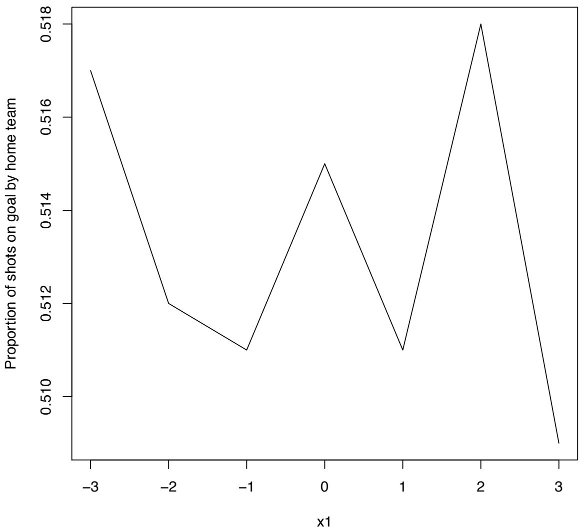
Is this because:

1. the game situation causes teams to play differently?
2. the game situation causes referees to officiate differently?

Change in Playing Style due to Goal Differential?



Change in Playing Style due to Penalty Differential?



Temporal Changes

Original Logistic Regression:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.124 + 0.401x_1 - 0.052x_2 - 0.030x_4 - 0.005x_1x_3$$

Logistic Regression from First Half of Games:

$$\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = -0.113 + 0.489x_1 - 0.047x_2 - 0.043x_4 - 0.009x_1x_3$$

Assessing an Aspect of Officiating Bias

- a simpler model: $\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}_0 + \hat{\beta}_1 x_1 \quad \rightarrow \hat{p}(x_1)$
- excluding referee j: $\log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \hat{\beta}_0^{(j)} + \hat{\beta}_1^{(j)} x_1 \quad \rightarrow \hat{p}^{(j)}(x_1)$
- no referring bias: $\log\left(\frac{p}{1-p}\right) = 0 \quad \rightarrow p = 0.5$

Referee j is making worse than average decisions under x_1 if

$$|\hat{p}^{(j)}(x_1) - 0.5| < |\hat{p}(x_1) - 0.5| \ .$$

Assessing an Aspect of Officiating Bias

Previous development leads to the performance metric

$$Q_j = \sum_{x_1} w(x_1) \left(\, |\hat{p}^{(j)}(x_1) - 0.5| - |\hat{p}(x_1) - 0.5| \, \right)$$

where the weight $w(x_1)$ is the proportion of penalties in the data set corresponding to x_1.

Rankings

Referee	Matches	Measure $1000Q_j$	Standard Deviation
01. Peel, Tim	341	1.26	0.32
02. Walsh, Ian	343	0.96	0.30
03. Devorski, Paul	336	0.57	0.26
04. Pochmara, Brian	343	0.51	0.31
05. Dwyer, Gord	329	0.36	0.29
06. Morton, Dean	329	0.36	0.27
07. Kozari, Steve	345	0.34	0.34
08. O'Rourke, Dan	343	0.18	0.30
09. Lee, Chris	343	0.17	0.28
10. St. Pierre, Justin	318	0.10	0.23

Table: Performance measures and standard deviations for the top 10 referees with at least 300 games officiated during the 2009/2010 through 2013/2014 NHL regular seasons.

Thank-you and don't forget:

Vancouver Hockey Analytics Conference - April 9, 2016

www.stat.sfu.ca/hockey.html