Do Winners Do It Differently?: The In-Game Production of Winning and Losing Hockey Games

Weller Ross and Kevin Mongeon

Graduate Student and Assistant Professor of Sport Management at Brock University

June 4, 2015

Presentation Outline

Do high and low performing NHL teams play differently?

1. Motivation
2. Previous literature
3. Current research
4. In-game production of winning model
5. Data, empirical approach, and results
6. Sports analytics application
7. Concluding remarks

Motivation

In-game production of winning

- New model based on Bayesian updating.

Sports analytics application

- Model results are used to discern information relating to team-specific behaviours in terms of responding to goal scored and allowed incentives.

Previous Research

In-game production of winning

- Poisson and Brownian motion process (Washburn, 1991; Stern, 1994).
- Bivariate Poisson regression (Dimitris and Ntzoufras, 2003)
- Markov Model (Kaplan, Mongeon, and Ryan, 2013).

Applications

- Examine betting market efficiency (Choi and Hui, 2012; Croxson and Reade, 2014).
- In-game strategies (Morrison, 1976).
- Decision-making (Abrevaya and McCulloch, 2004; Mongeon and Mittelhammer, 2013).

Current Research

Bayesian updating model to calculate in-game win probabilities throughout the progression of games.

- Deterministic formula used to obtain conditional probabilities.
- Based on established probability axioms.
- Win probabilities are obtained through straight-forward statistical calculations.

Optimal team behaviour

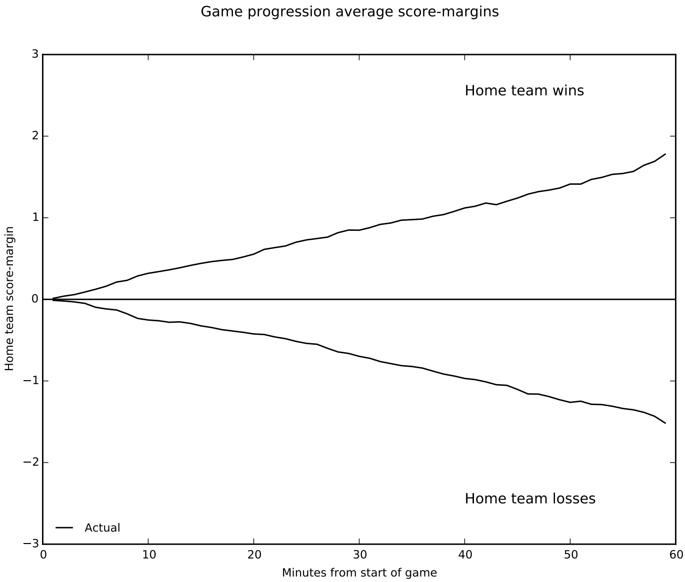
- High-performing team's manage in-game goal scored and allowed incentives better than low-performing teams.
- Case study: New York Rangers, Montreal Canadians, Toronto Maple Leafs, and Buffalo Sabres.
- Examine from the home team perspective.

In-Game Production of Winning Model

$$p(H_g|S_{g,t}) = \frac{p(S_{g,t}|H_g)p(H_g)}{p(S_{g,t}|H_g)p(H_g)+p(S_{g,t}|\bar{H}_g)p(\bar{H}_g)}$$

$$\frac{p(H_g|S_{g,t})}{p(H_g|\bar{S}_{g,t})} = \frac{p(H_g)}{p(\bar{H}_g)} \times \frac{p(S_{g,t}|H_g)}{p(S_{g,t}|\bar{H}_g)}$$

posterior probability = prior odds $\times$ likelihood ratio



Game progression average score-margins
Home team wins
Home team losses
Home team score-margin
Minutes from start of game
Actual
3
2
1
0
−1
−2
−3
0
10
20
30
40
50
60

Likelihood Ratio Estimation Procedure

Empirically describe the underlying game-state process separately across games won and lost as a generalized Poisson process

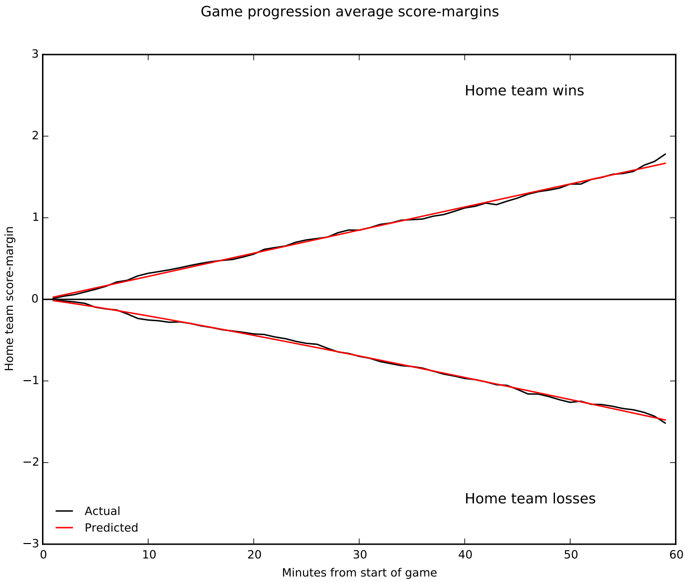
$$y_{g,t} = (y_{h,k,t}, y_{a,k,t} | H_k, \bar{H}_k) \sim Poisson(\lambda_{h,k,t}, \lambda_{a,k,t})$$

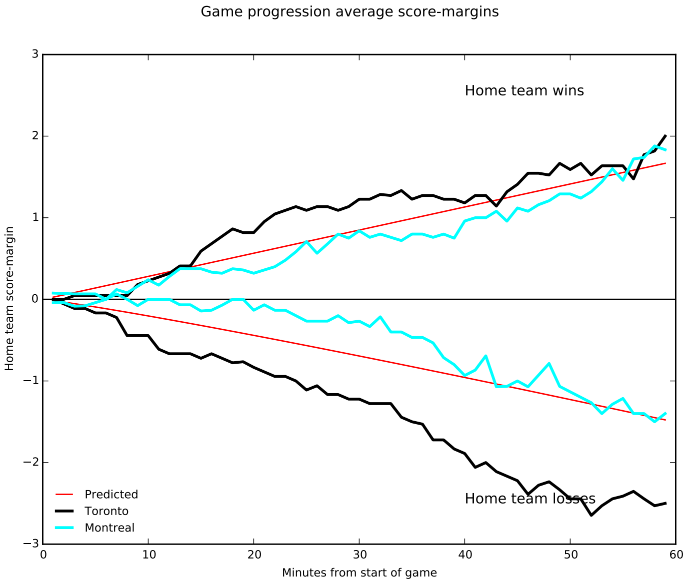
$$log(\lambda_{h,k,t}) = \beta_0 + \beta_1 minRem_{k,t} + \epsilon_{h,k,t}$$

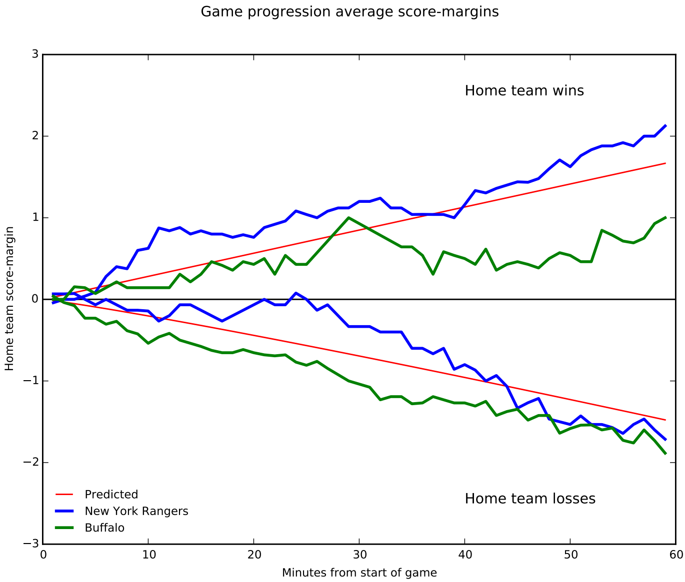
$$log(\lambda_{v,k,t}) = \theta_0 + \theta_1 minRem_{k,t} + \epsilon_{a,k,t}$$

for historical games played during the 2013 season, where $k = 1, ..., 1230$, and $t = 1, ..., 60$.

$$E(SM_{k,t} | H_k, \bar{H}_k) = E(\lambda_{h,k,t}, \beta) - E(\lambda_{a,k,t}, \theta)$$


Game progression average score-margins
Home team wins
Home team losses
Home team score-margin
Minutes from start of game
Actual
Predicted
3
2
1
0
−1
−2
−3
0
10
20
30
40
50
60


Game progression average score-margins
Home team wins
Home team losses
Home team score-margin
Minutes from start of game
Predicted
Toronto
Montreal
3
2
1
0
−1
−2
−3
0
10
20
30
40
50
60



Game progression average score-margins
Home team wins
Home team losses
Home team score-margin
Minutes from start of game
3
2
1
0
−1
−2
−3
0
10
20
30
40
50
60
Predicted
New York Rangers
Buffalo

Game progression posterior probabilties

For 2014 games,

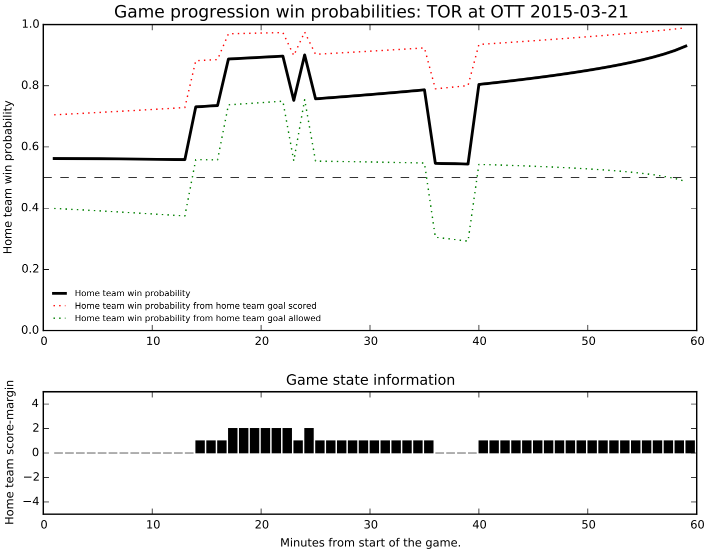
$$\frac{p(H_g|S_{g,t})}{p(H_g|\bar{S}_{g,t})} = \frac{p(H_g)}{p(\bar{H}_g)} \times \frac{p(S_{g,t}|H_g)}{p(S_{g,t}|\bar{H}_g)}$$

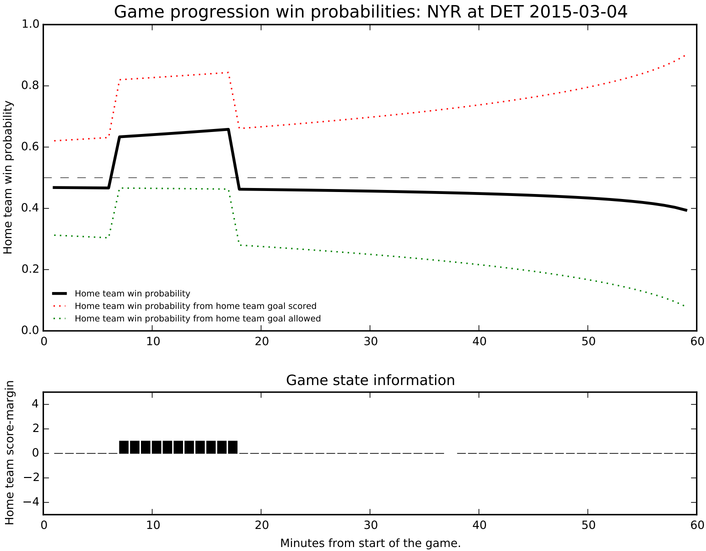
$$p(H_g) = \frac{YTDHomeWP_g}{YTDHomeWP_g + YTDAwayWP_g}$$

$$p(\bar{H}_g) = 1 - p(H_g)$$

$$p(S_{g,t}|H_g) = \frac{e^{\lambda_{h,k,t}}, \lambda_{h,k,t}^{y_{h,k,t}}}{y_{h,k,t}!} \frac{e^{\lambda_{a,k,t}} \lambda_{a,k,t}^{y_{a,k,t}}}{y_{a,k,t}!} |H_k$$

$$p(S_{g,t}|\bar{H}_g) = \frac{e^{\lambda_{h,k,t}}, \lambda_{h,k,t}^{y_{h,k,t}}}{y_{h,k,t}!} \frac{e^{\lambda_{a,k,t}} \lambda_{a,k,t}^{y_{a,k,t}}}{y_{a,k,t}!} |\bar{H}_k$$


Game progression win probabilities: TOR at OTT 2015-03-21
Home team win probability
Home team win probability
Home team win probability from home team goal scored
Home team win probability from home team goal allowed
Game state information
Home team score-margin
Minutes from start of the game.



Game progression win probabilities: NYR at DET 2015-03-04
Home team win probability
Home team win probability
Home team win probability from home team goal scored
Home team win probability from home team goal allowed
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
Game state information
Home team score-margin
4
2
0
−2
−4
Minutes from start of the game.

Average Team-Specific In-Game Win Probabilties

Team	All states	1^{st} Intermission	2^{nd} Intermission
Wins			
Buffalo	0.48	0.42	0.50
Toronto	0.56	0.58	0.59
Montreal	0.67	0.66	0.68
New York	0.79	0.68	0.75
Losses			
Buffalo	0.24	0.29	0.20
Toronto	0.27	0.35	0.22
Montreal	0.54	0.65	0.50
New York	0.41	0.57	0.36

Sport Analytics Application

Optimal team behaviour

- A trade-off exists between playing offense and defense within teams.
- Asymmetric offensive and defensive incentives exist across competing teams.

Win probability state:	greater than 0.50	less than 0.50
Average benefit of a goal scored	0.13	0.20
Average cost of a goal allowed	0.19	0.14
Net benefit	-0.06	0.06

Sport Analytics Application

Optimal team behaviour

- Increase total production as the difference in the benefit of goal scored and cost of a goal allowed increases.
- $\frac{\partial(production)}{\partial(benefit-cost)} > 0$

Hypothesis

- High-performing teams respond to goal scored and allowed incentives better than low-performing teams.

Sport Analytics Application

Optimal team behaviour

$$prod_{g,t} = \alpha + X'_{g,t}\mathrm{B} + \theta Z_{g,t} + \epsilon_{g,t}$$

- $prod$: goal score or goal allowed per 60 minutes.
- X: team indicator variable and the difference in the benefit of a goal score and the cost of a goal allowed interaction.
- Z: prior probability of a home team win.
- Unit of observation: game-minute increments.

The impact of net benefits/costs of goals on production.

Variable	Mean	Standard error
Buffalo	-2.51	(5.22)
Toronto	-2.49	(6.41)
Montreal	0.84	(5.63)
New York	2.00	(5.09)
Prior probability	-2.55	(2.44)
Constant	5.90	(1.26)

Concluding Remarks

- Bayes' Theorem to calculate in-game with probabilities throughout the progression of games.
- Application of the model results suggest that high-performing teams respond to goal incentives better than low-performing teams.
- Low-performing teams can potentially improve without roster adjustments.
- Future research relating stochastic frontier models to estimate team-specific in-game production of winning efficiencies.

Thank you.