Scoring Trends in the Third Period in the NHL

Jack Davis
Simon Fraser University

Joint work with Dr. Paramjit Gill,
University of British Columbia - Okanagan

Three eras based on regular season payoff structure:

Three eras based on regular season payoff structure:

- Zero-Sum Era (Up to 1998-99)
  5 minute OT
  2pts for a win, 1pt for a tie. 0 for a loss

Three eras based on regular season payoff structure:

- Zero-Sum Era (Up to 1998-99)
  5 minute OT
  2pts for a win, 1pt for a tie. 0 for a loss
- OT Loss Era (Up 1999-2000 to 2003-04)
  5 minute OT
  2pts for a win, 1pt for a tie, 1pt for an OT loss.

Three eras based on regular season payoff structure:

- Zero-Sum Era (Up to 1998-99)
  5 minute OT
  2pts for a win, 1pt for a tie. 0 for a loss
- OT Loss Era (Up 1999-2000 to 2003-04)
  5 minute OT
  2pts for a win, 1pt for a tie, 1pt for an OT loss.
- Shootout Era (2005-06 until now)
  5 minute OT followed by a shootout
  2pts for a win, 1pt for an OT or shootout loss.

We ask two questions:

- 1. Has the scoring rate in the 3rd period conditional on situation changed?

We ask two questions:

- 1. Has the scoring rate in the 3rd period conditional on situation changed?
- 2. Has the probability of a game going to overtime changed?

N = Number of intervals sampled. Truncation at $t^* = 300$.

m = Times out of N an interval ends early because of a goal.

t_i = Length (seconds) of i^{th} short interval, i=1,...,m

$\hat{\theta}$ = MLE estimate of time to goal (Bartholomew 1957)

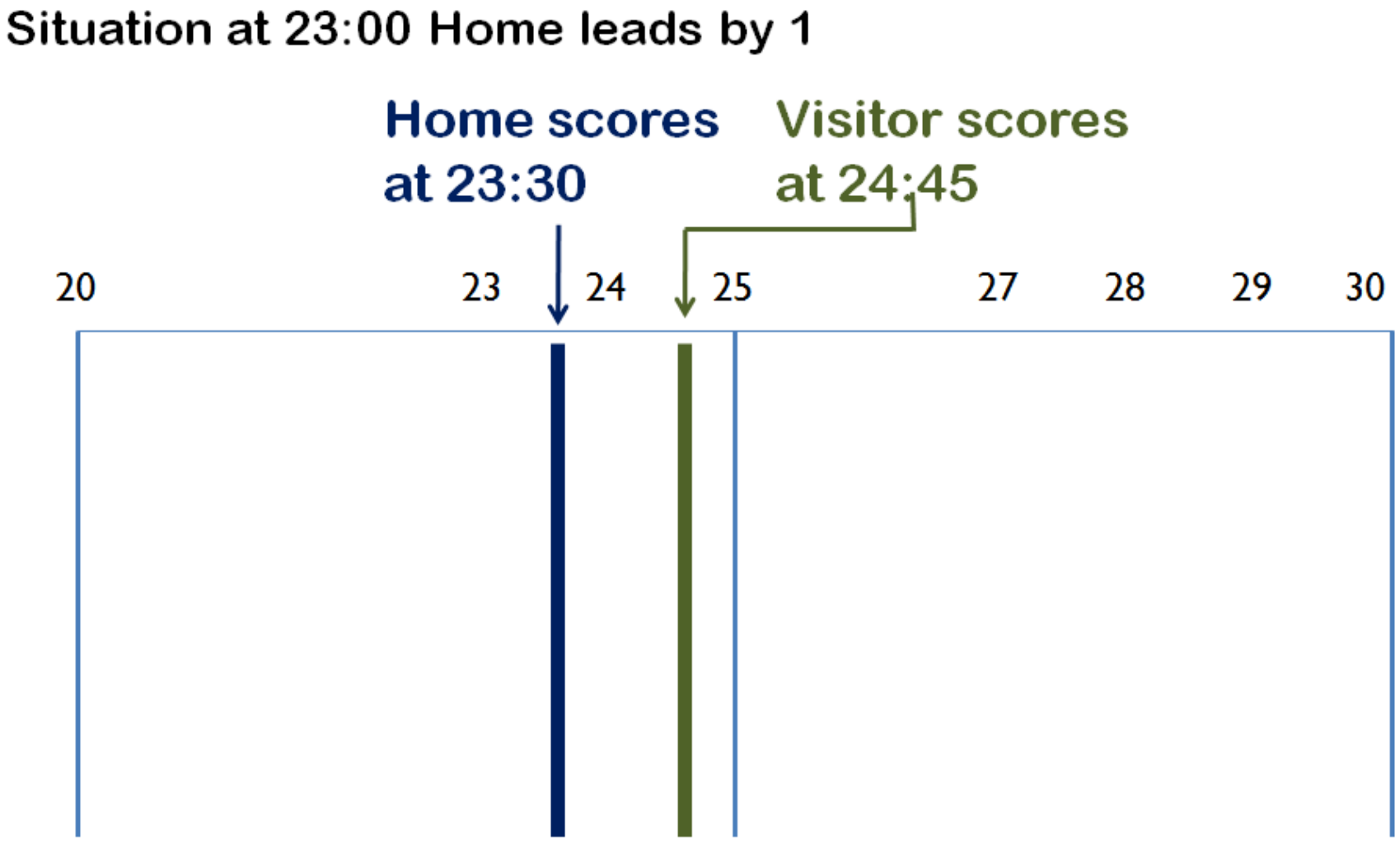
$$\hat{\theta} = \frac{\sum t_i + (N - m)t^*}{m}$$

Scoring intensity $= 3600/\hat{\theta}$ = Goals per hour

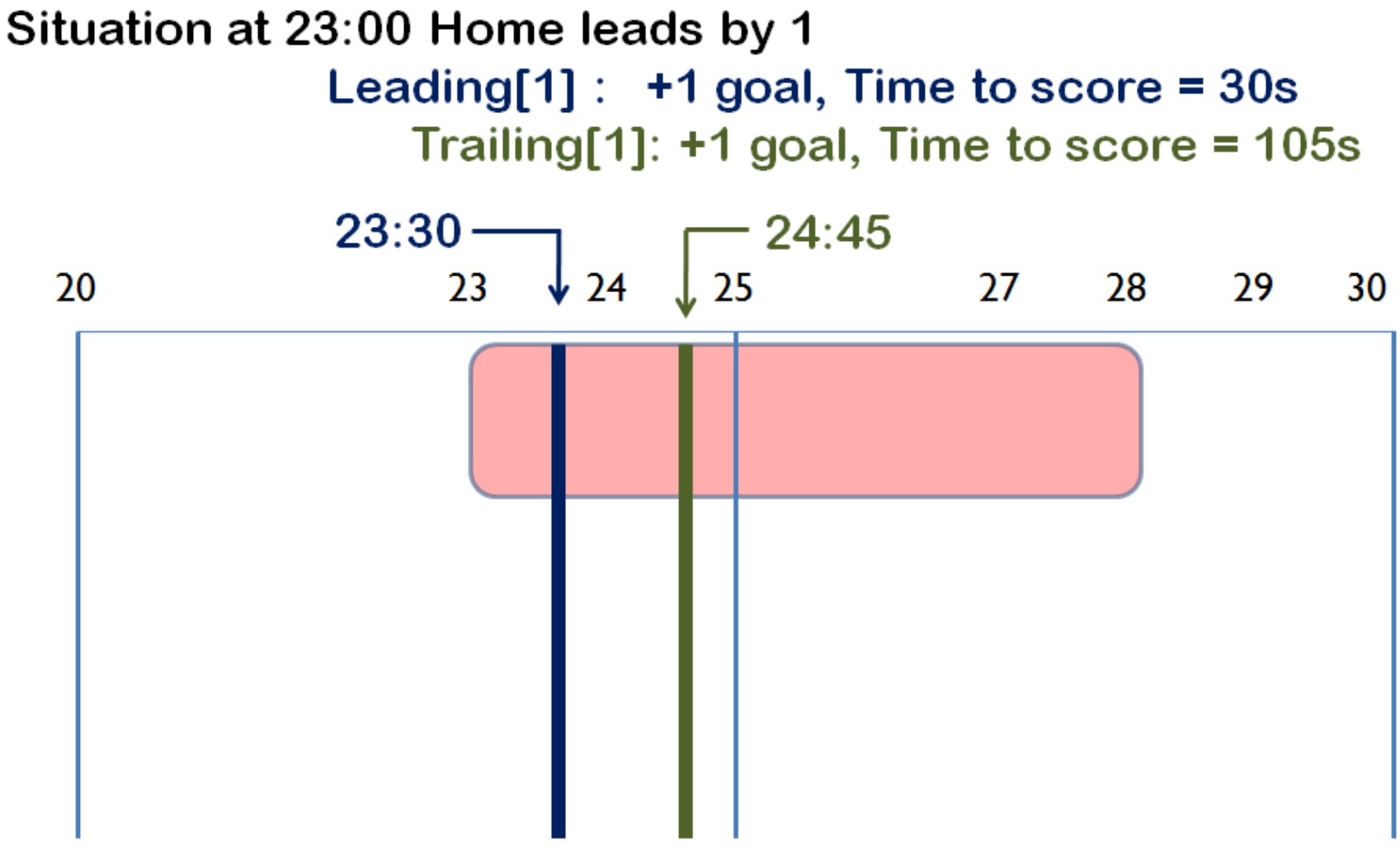
Truncated exponential model applied to find the goal rate at each minute of regulation play.

Each situation (ahead by 1, tied, behind by 1, etc.) is considered seperately and measured at the *beginning* of the interval.

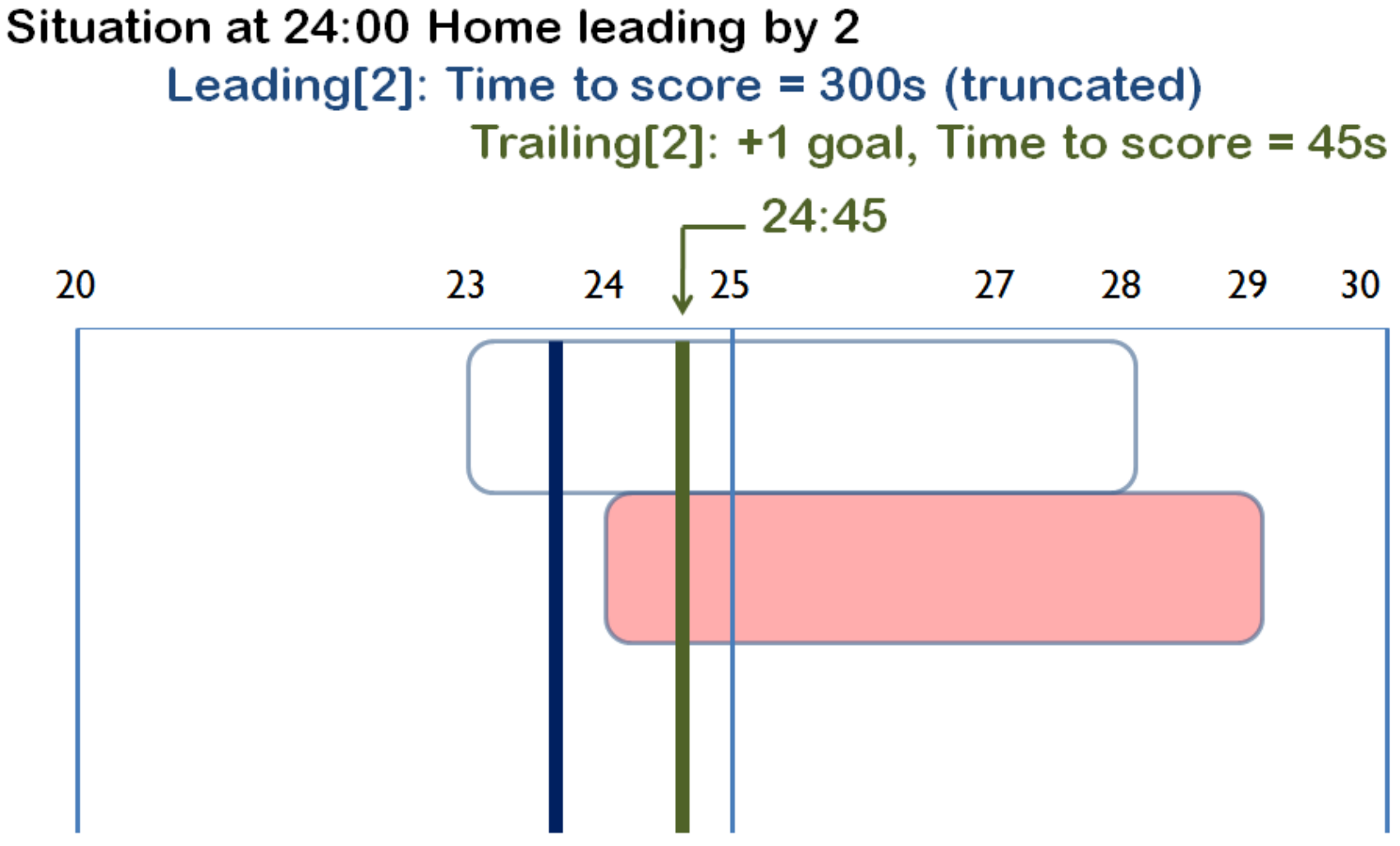
Rolling truncated example.



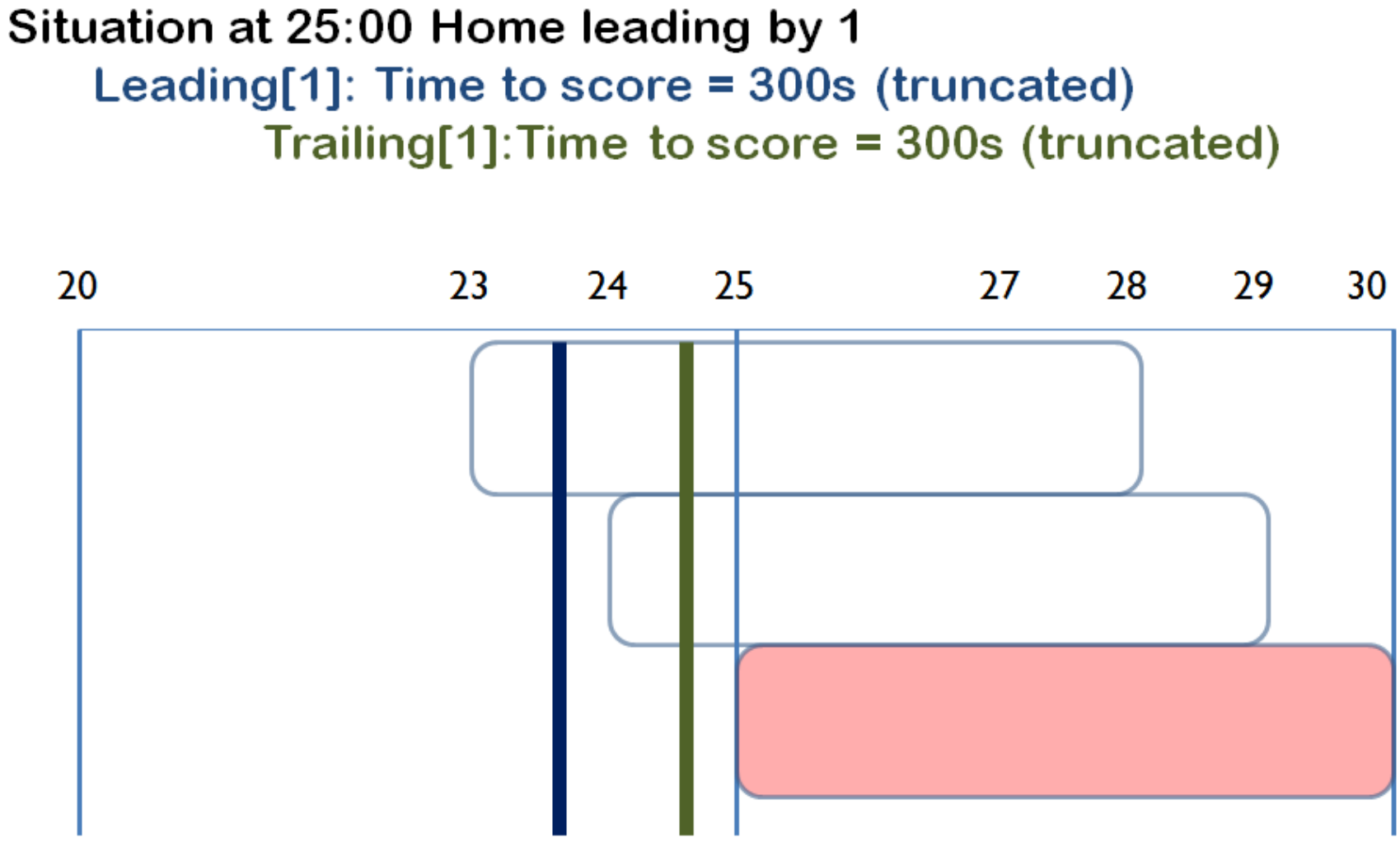
Rolling truncated example.



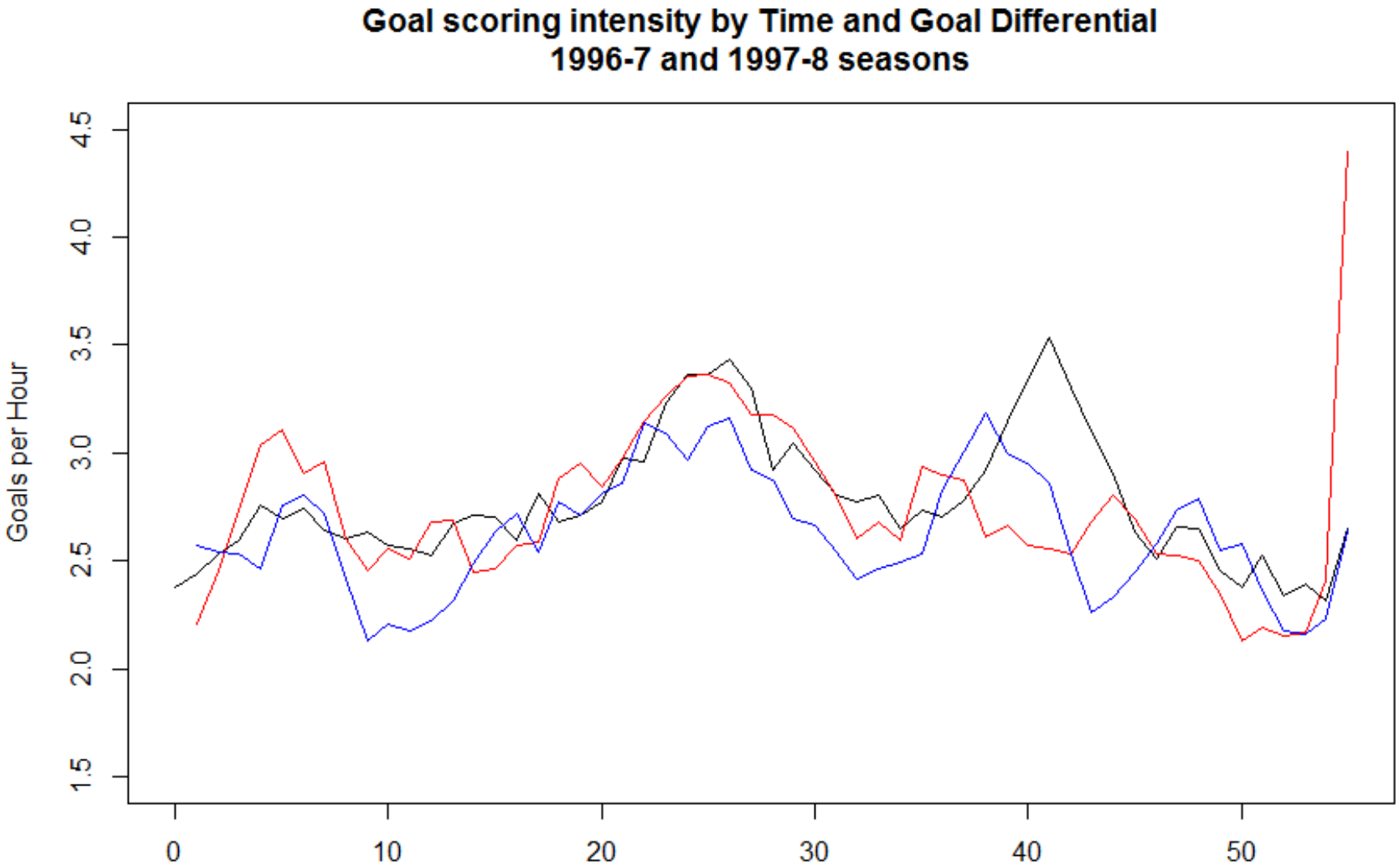
Rolling truncated example.



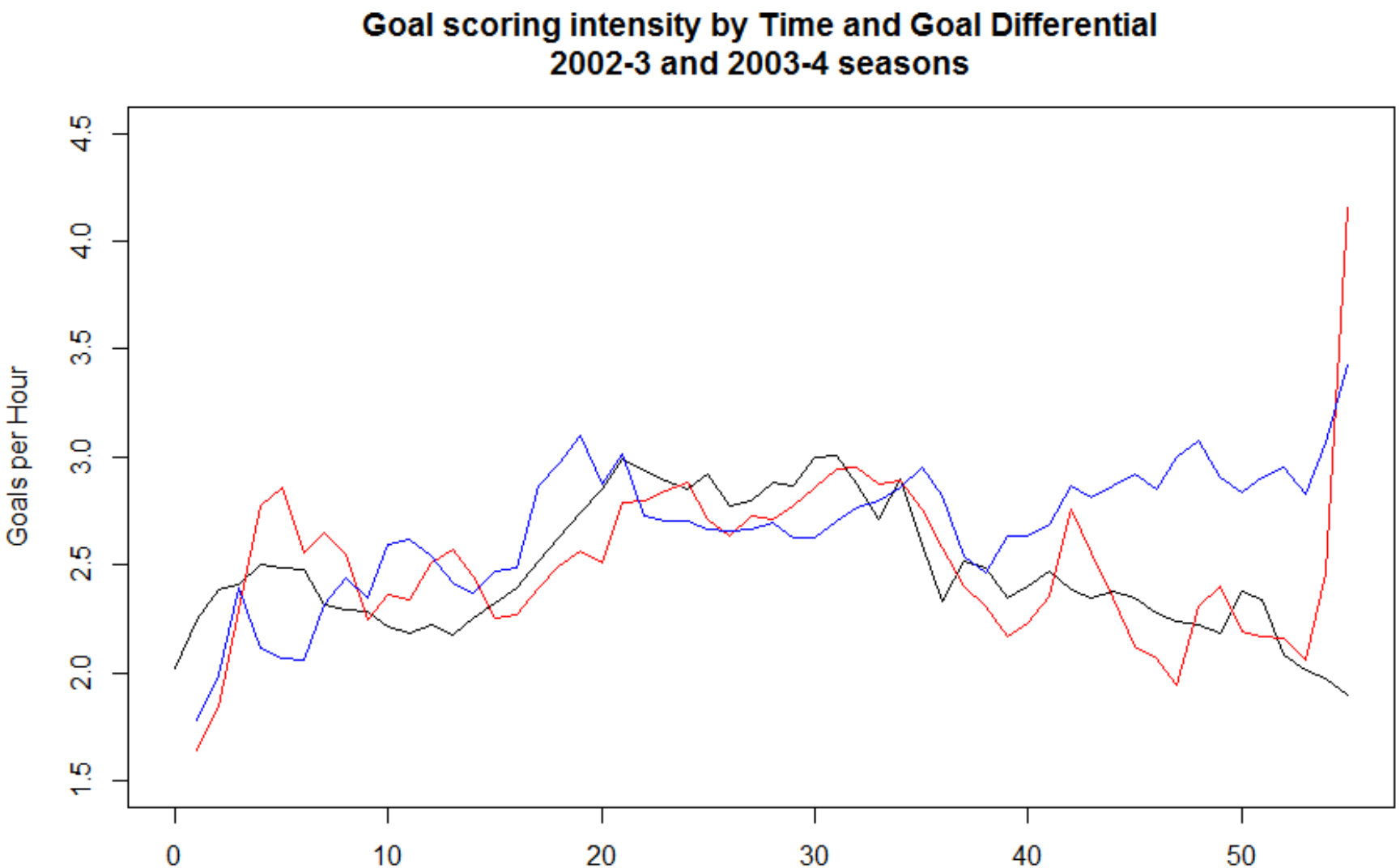
Rolling truncated example.



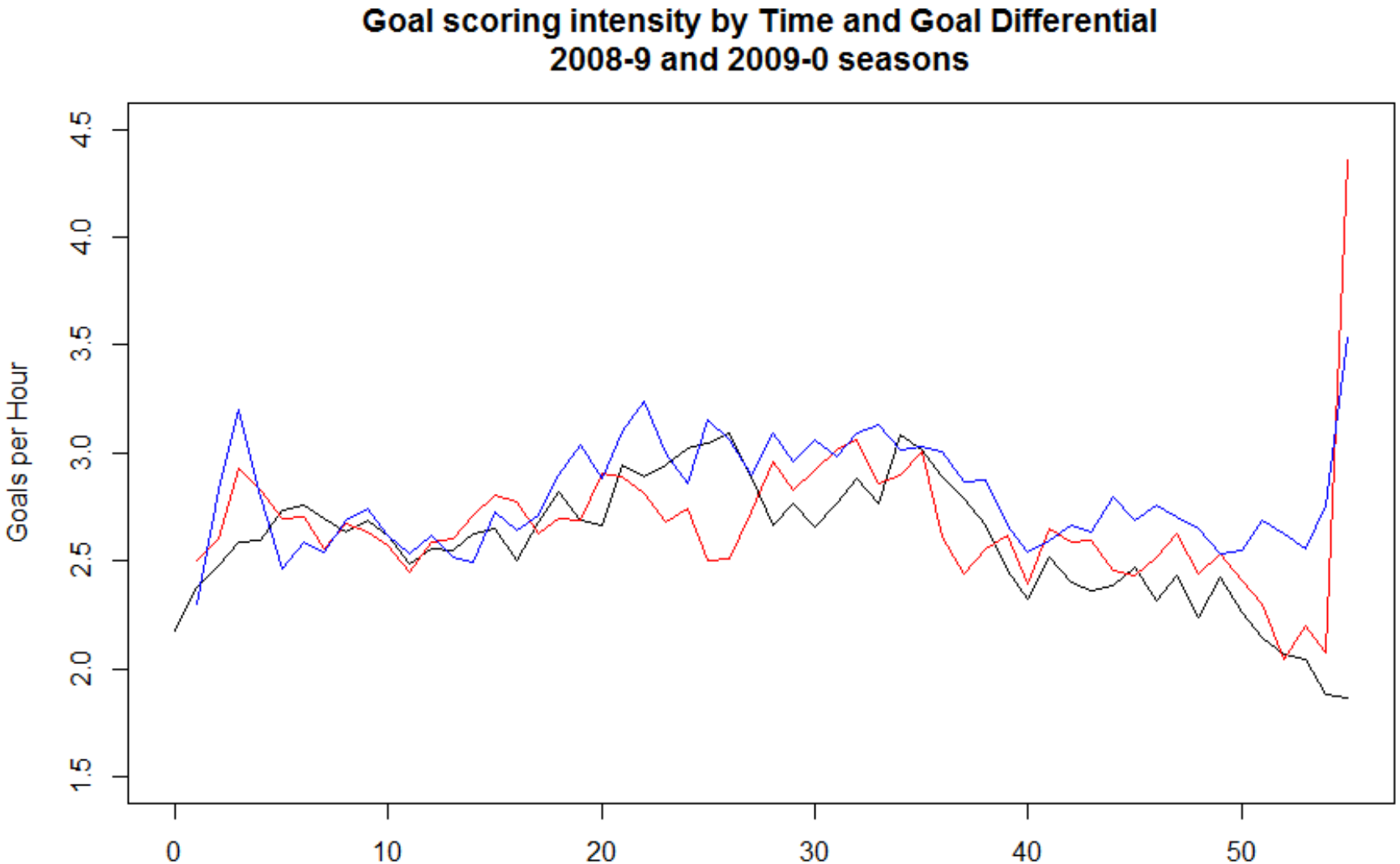
Scoring Intensity for 1997-98 and 1998-99 seasons (Zero-sum era)



Scoring Intensity for 2002-3 and 2003-4 seasons (OT-Loss era)



Scoring Intensity for 2007-08 and 2008-09 seasons (Shootout era)

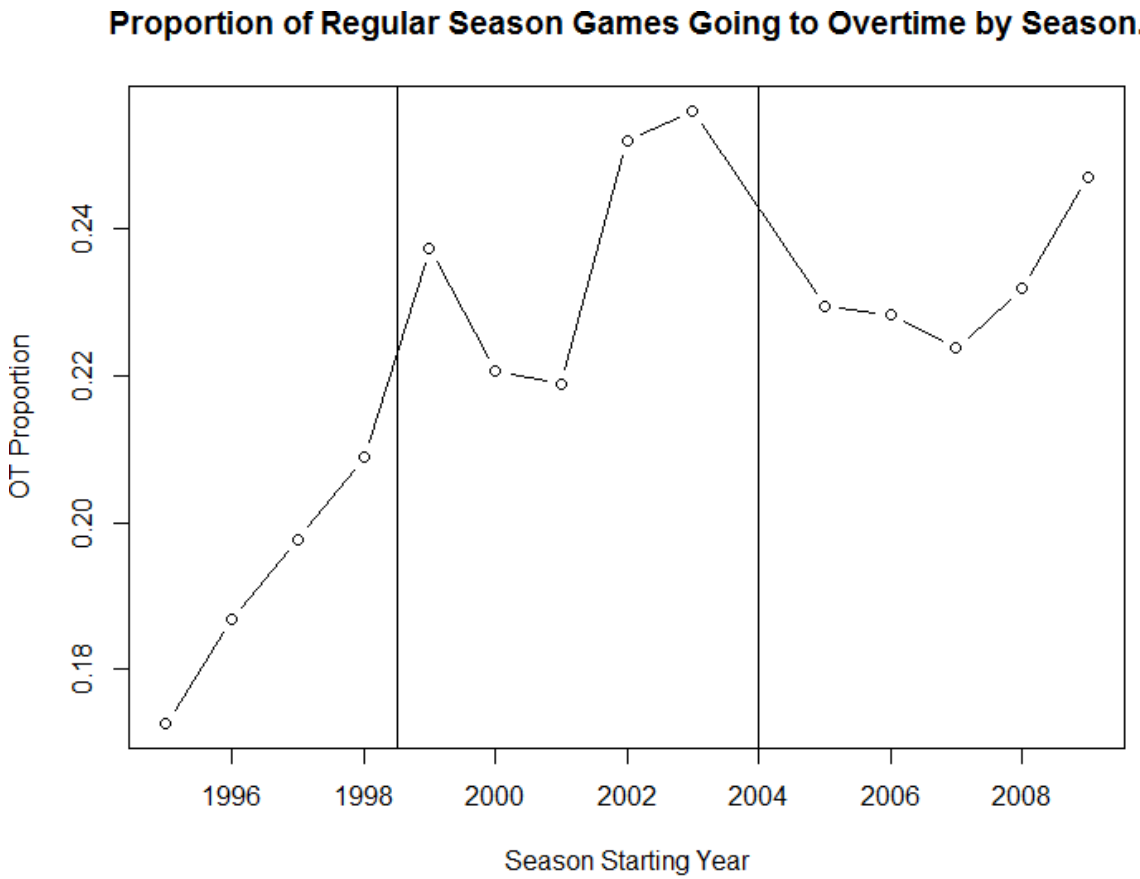


All eras have a spike in goal intensity in the last five-minute block, especially in the leading intensity. (Empty net?)

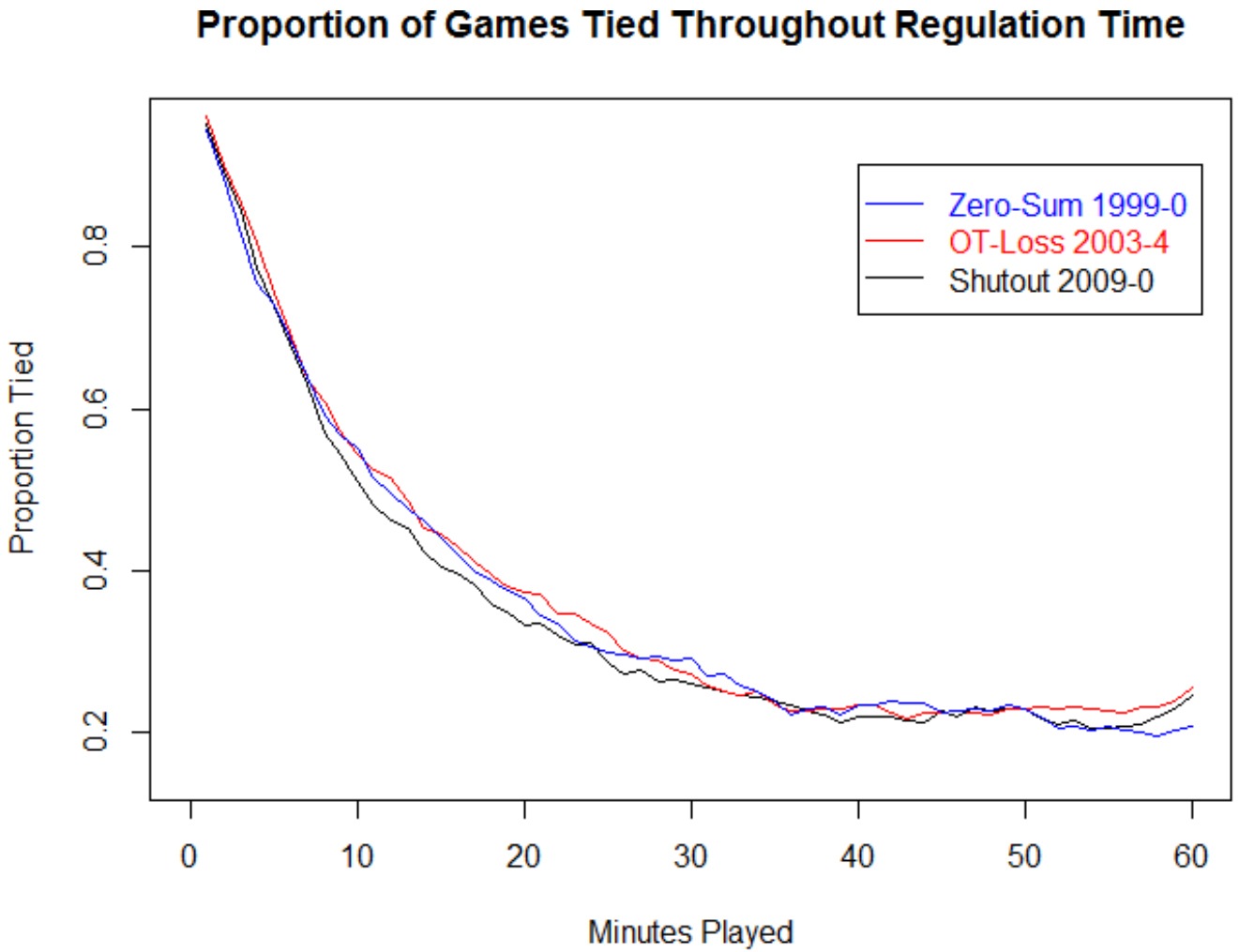
After the OT Loss rule, a divergence appears between the tied situation and other situations.

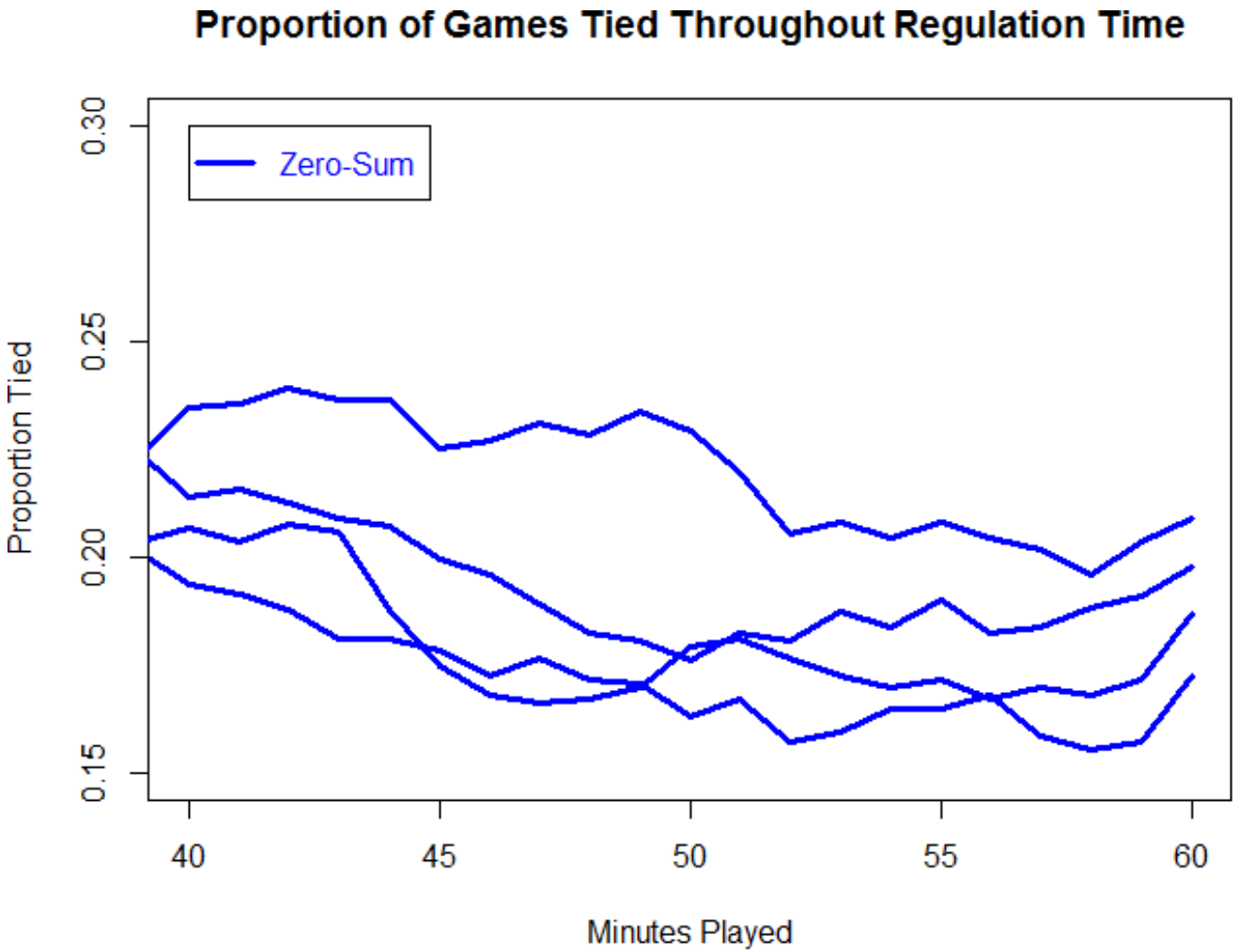
The divergence is less pronounced but starts near the same time in the Shootout Era.

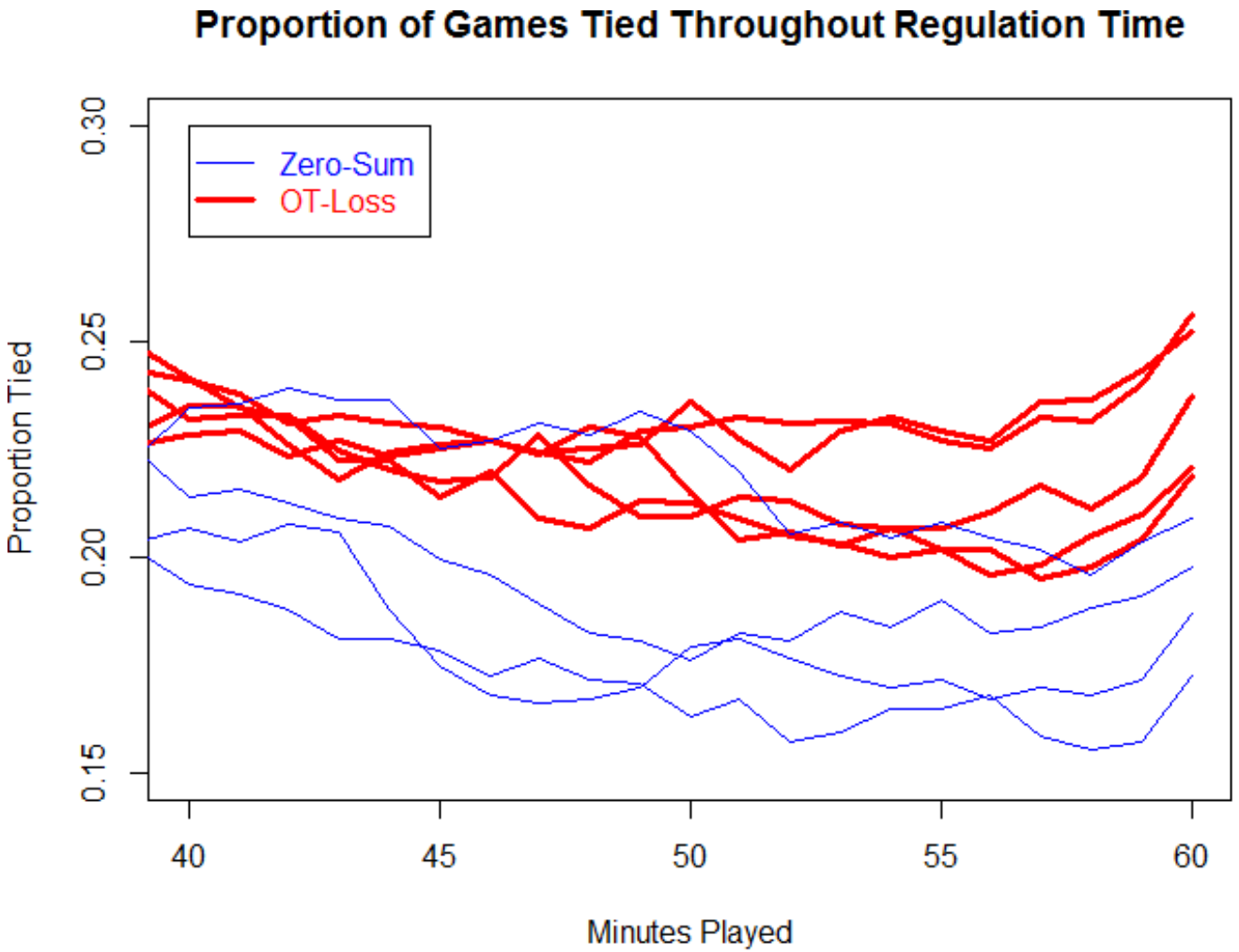
Proportion of games in OT by season

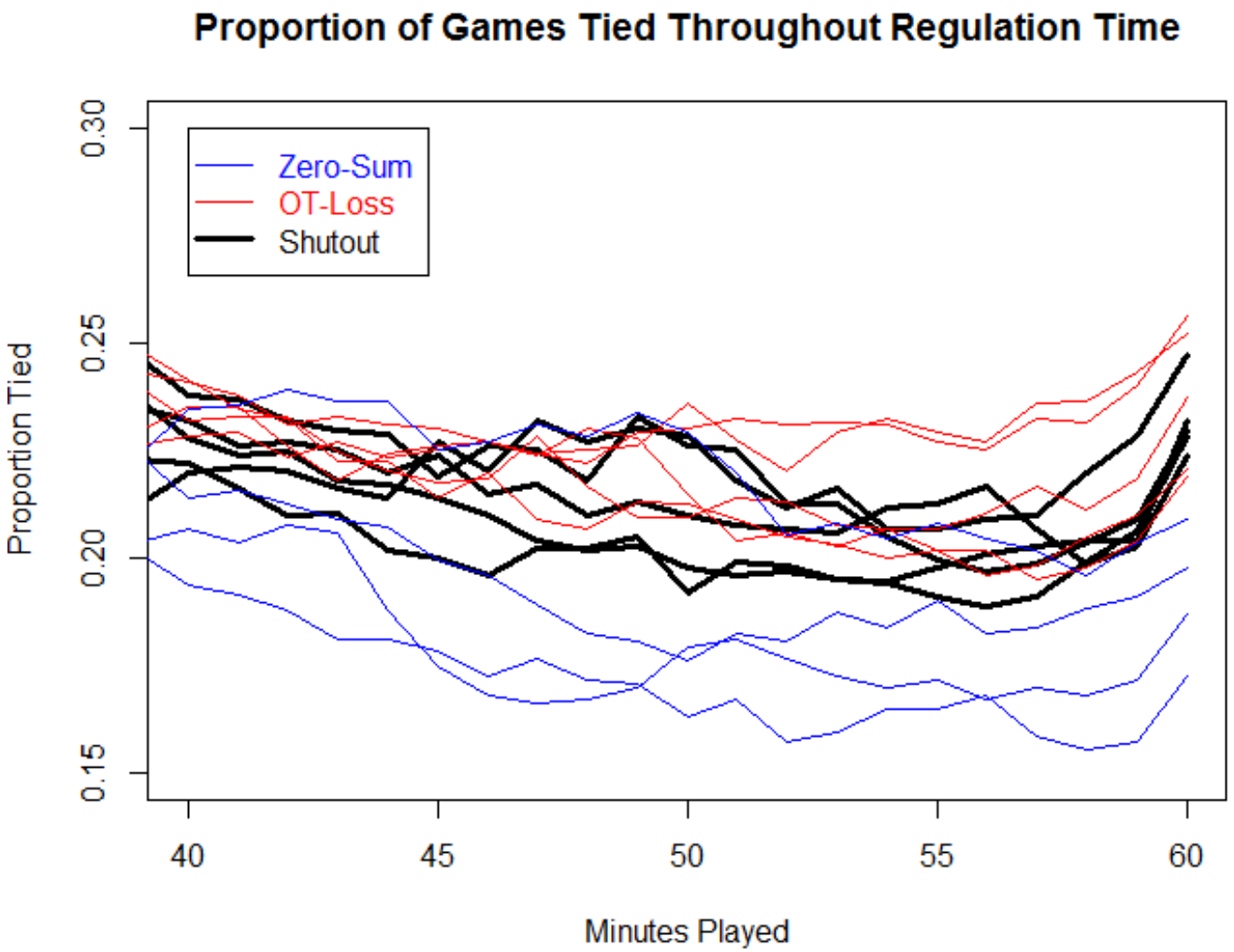


Proportion of games tied after X minutes




Proportion of Games Tied Throughout Regulation Time
Zero-Sum
Proportion Tied
0.15
0.20
0.25
0.30
40
45
50
55
60
Minutes Played


Proportion of Games Tied Throughout Regulation Time
Zero-Sum
OT-Loss
Proportion Tied
0.15
0.20
0.25
0.30
40
45
50
55
60
Minutes Played


Proportion of Games Tied Throughout Regulation Time
Zero-Sum
OT-Loss
Shutout
Proportion Tied
0.15
0.20
0.25
0.30
40
45
50
55
60
Minutes Played

A Poisson family generalized linear model (GLM) is fit to the number of goals scored in regulation play by a team.

The scoring team's overall offensive ability (average goals scored), the opposing team's defensive ability, and an indicator of that the scoring team is at home were used as predictors.

λ_{ijk} is the expected number of goals scored during regulation play of a game by team i against team j during home-game situation k.

$$\log(\lambda_{ijk}) = \alpha + Off_i + Def_j + HomeTeam_k\phi$$

$$\log(\lambda_{ijk}) = \alpha + Off_i + Def_j + HomeTeam_k\phi$$

- α is the baseline scoring rate
- Off_i is the offensive skill of the team were modeling.
- Def_j is the defensive skill of the team opposing the one were modeling.
- ϕ is the league home team advantage, which is only included for the home team.

Karlis & Ntzoufras (2003) mixture model x = Home team score; y = Away team score, where

$$Pr(x, y) =$$
$$(1 - p) \times Pois(x, \lambda_H) \times Pois(y, \lambda_A) \quad \text{when x} \neq \text{y}$$
$$(1 - p) \times Pois(x, \lambda_H) \times Pois(y, \lambda_A) + pD(x, \theta) \quad \text{when x} = \text{y}$$

Diagonal (tie) inflation proportion= p
$D(x, q) = Poisson(q)$

$$Pr(x, y) =$$
$$(1-p) \times Pois(x, \lambda_H) \times Pois(y, \lambda_A) \quad \text{when x} \neq \text{y}$$
$$(1-p) \times Pois(x, \lambda_H) \times Pois(y, \lambda_A) + pD(x, \theta) \quad \text{when x} = \text{y}$$

When modelling both of the scores in a game, the standard Poisson model is the inflated model when $p = 0$.

A likelihood ratio test indicates that the diagonally inflated model explains the data better.

In OT loss; shootout eras, diagonal inflation improves the model. The structural diagonal proportion p in DIDP model) is larger in OT loss; shootout eras.

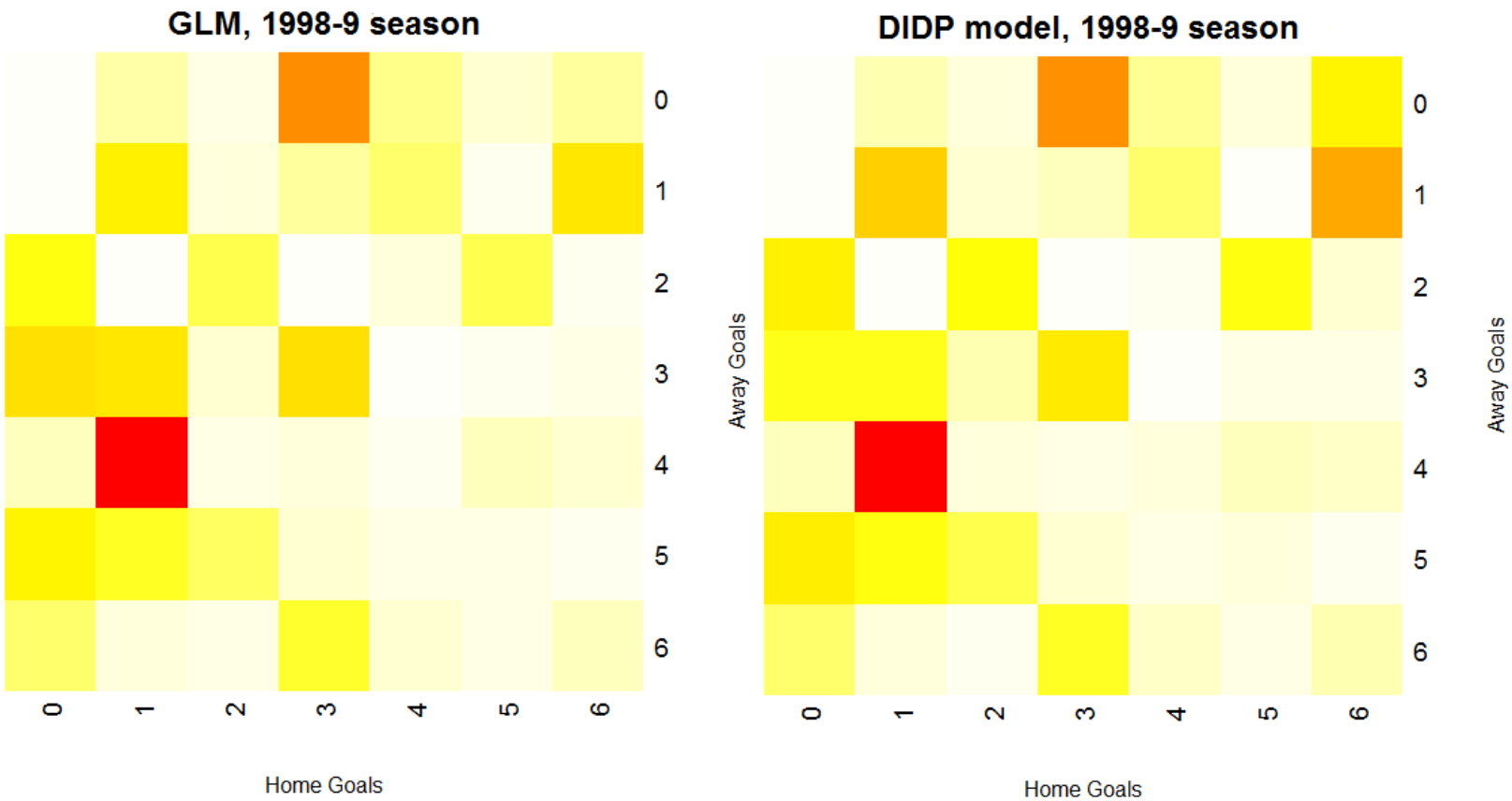
Season	DIDP mixture p	Liklihood Ratio Test p-value
1996-7	.005	.0874
1997-8	.005	.0918
1998-9	$<$.001	1
2001-2	.047	.00011
2002-3	.074	$<$.00001
2003-4	.088	$<$.00001
2007-8	.044	.0011
2008-9	.069	$<$.00001
2009-0	.075	$<$.00001

χ^2_{36} stats of tables of expected vs. actual home-away outcomes. Ignores games where a team scores 7+ goals.

Season	GLM χ^2_{36}	DIDP χ^2_{36}
1996-7	70.1	66.5
1997-8	78.8	72.7
1998-9	58.7	58.5
2001-2	98.1	86.3
2002-3	112.2	73.1
2003-4	112.6	100.5
2007-8	89.8	76.0
2008-9	102.1	77.7
2009-0	132.6	88.6

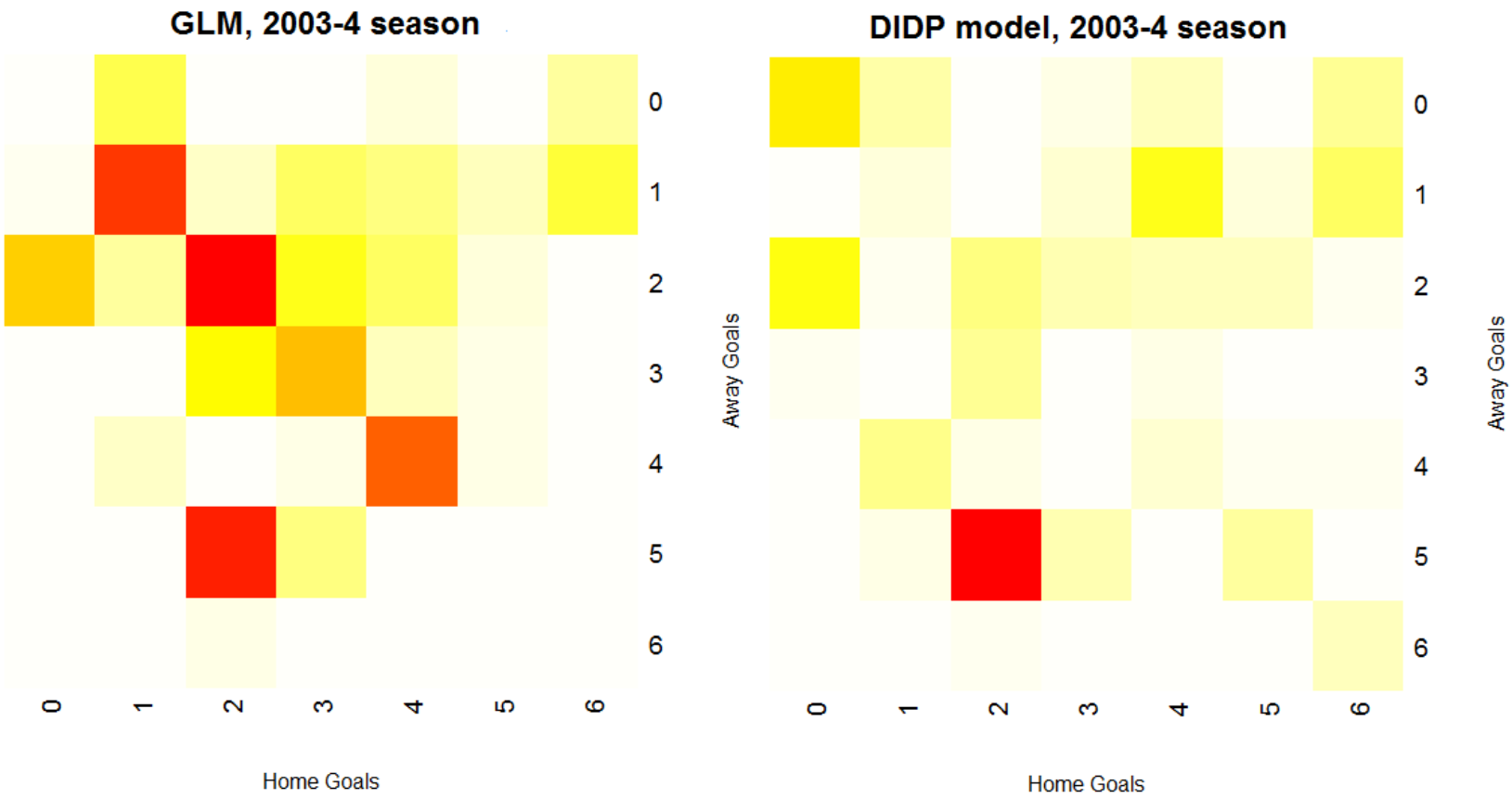
$$\chi^{2^*}_{36,.95} = 51.0,\ \chi^{2^*}_{36,.99} = 58.6$$

This heatmap shows the relative contribution to the χ^2 statistic for each season and model.



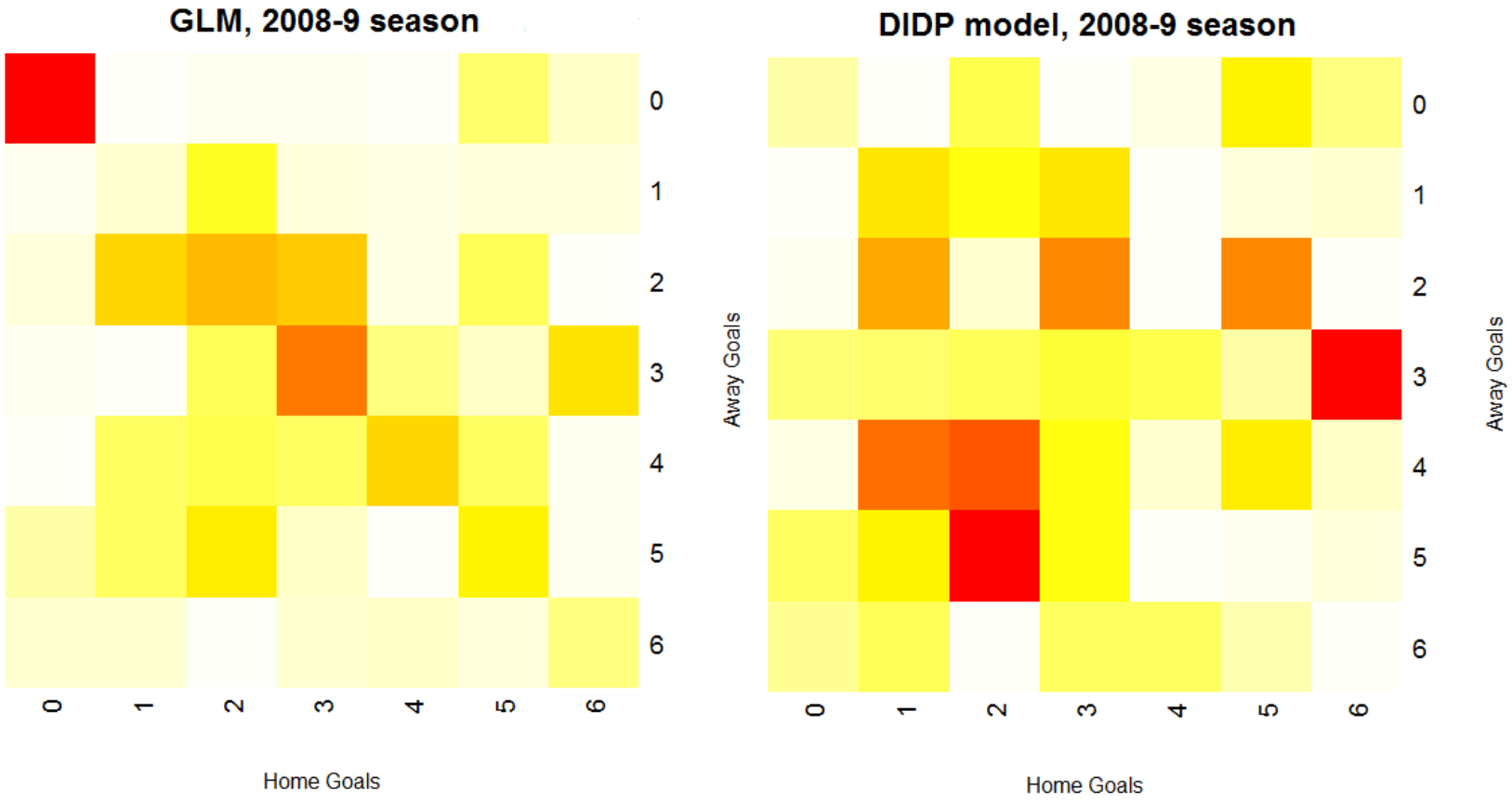
White squares indicate good fits to actual data and red squares poor.

This heatmap shows the relative contribution to the χ^2 statistic for each season and model.



White squares indicate good fits to actual data and red squares poor.

This heatmap shows the relative contribution to the χ^2 statistic for each season and model.



White squares indicate good fits to actual data and red squares poor.

In OT loss; shootout eras, diagonal inflation improves the model. The structural diagonal proportion p in DIDP model) is larger in OT loss; shootout eras.

Season	DIDP mixture p	Liklihood Ratio Test p-value
1996-7	$<.001$	1
1997-8	$<.001$	1
1998-9	.001	1
2001-2	.010	.0094
2002-3	.005	.0572
2003-4	.013	.0005
2007-8	.012	.0707
2008-9	.023	$\sim$.00001
2009-0	$<.001$	1