Predicting playoff outcomes based on regular season data

Nilesh Shah, PhD
Post-doctoral scholar
University of Pittsburgh

November 8, 2014

Introduction

- Can we predict postseason success based on regular season data?

- What data is useful?

- Is easily available data sufficient?

- Is hockey too random?

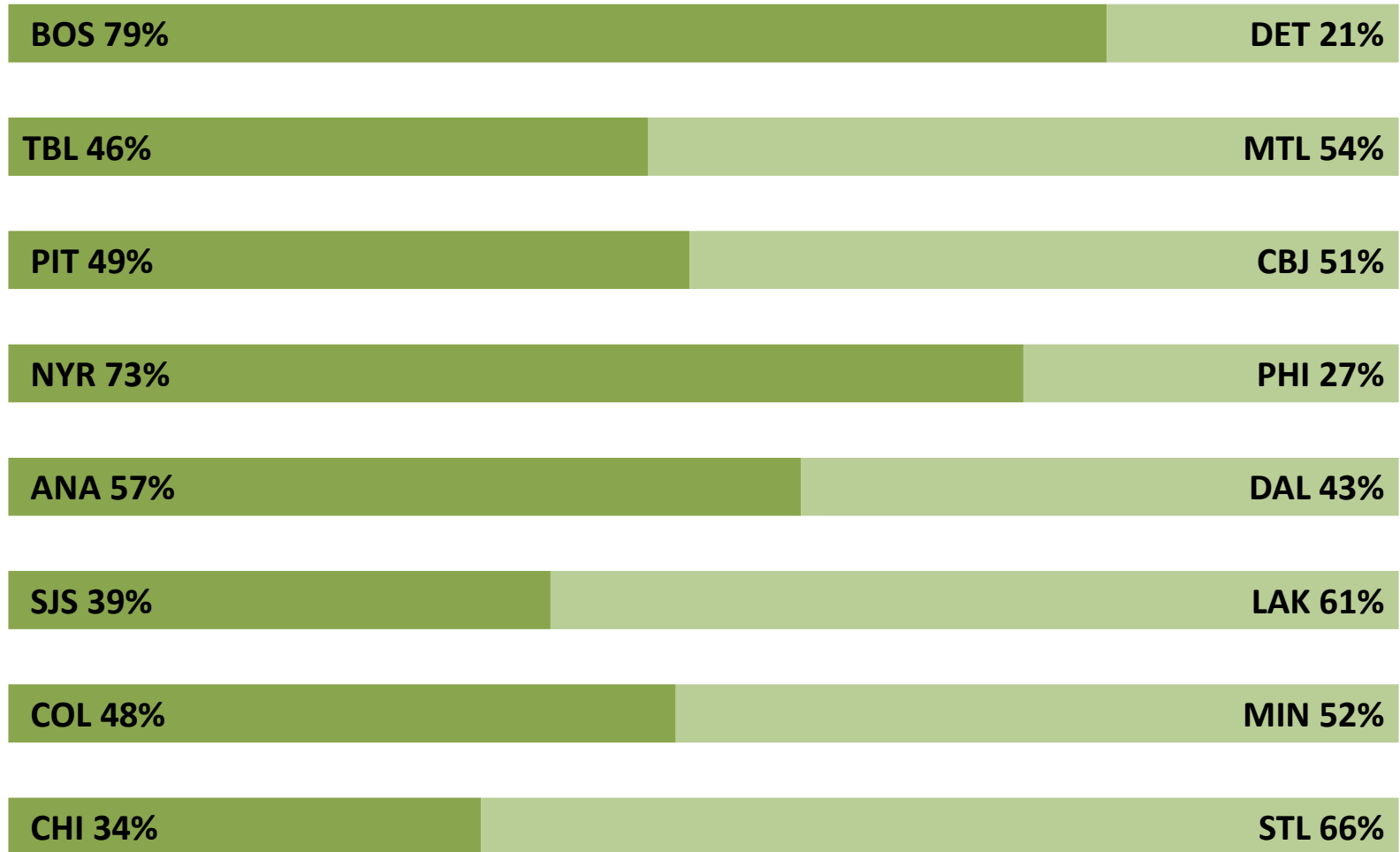- Can we at least get an idea of who's underrated or overrated?

Model considerations

- Outcome: home team series victory.
- Variables considered:
  - Score-adjusted Fenwick % (Tulsky)
  - Score-adjusted Corsi %
  - 5v5 save percentage
  - Power play %
  - Penalty kill %
  - Total points
  - Points in last 20 games
  - Seeding
- Sources: hockeyanalysis.com and espn.com

The model

- Data from 2007/2008-2012/13.
- n= 90 playoff series.
- Use logistic regression to model home team series win
- Backward stepwise regression used for variable selection.
- Final model included:
  - 5v5 Fenwick % difference
  - 5v5 team save percentage difference
  - Penalty kill % difference
- Model allows us to estimate the probability of a home team series win for each matchup.

First Round

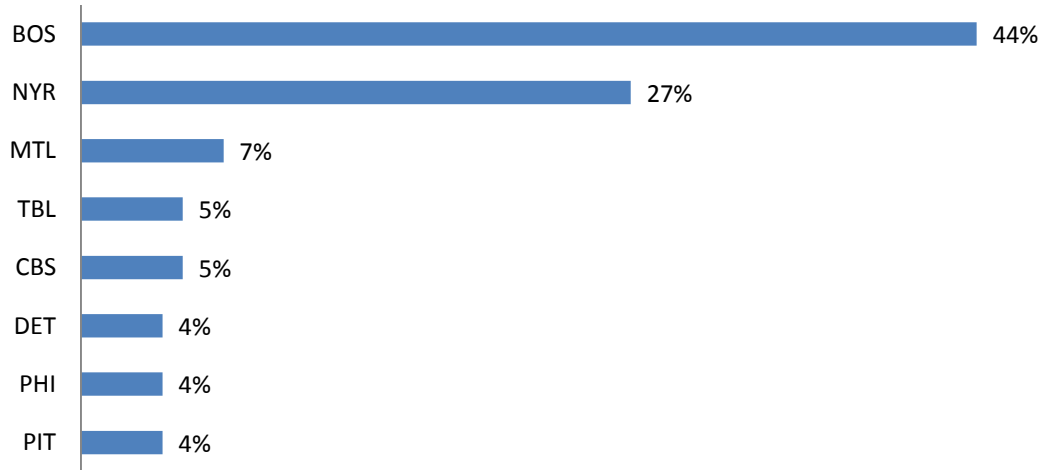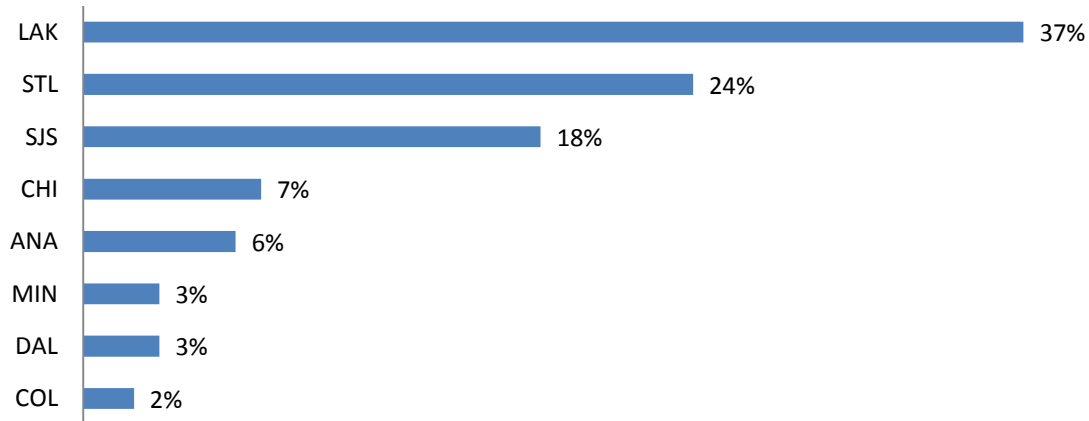|	
BOS 79%	**DET 21%**
TBL 46%	**MTL 54%**
PIT 49%	**CBJ 51%**
NYR 73%	**PHI 27%**
ANA 57%	**DAL 43%**
SJS 39%	**LAK 61%**
COL 48%	**MIN 52%**
CHI 34%	**STL 66%**

Simulations for 2014 playoffs

- Use series probability estimates to simulate how the playoffs will go.

- Create a random binomial variable based on the series probability.

- Simulate playoffs 10,000 times.

First round

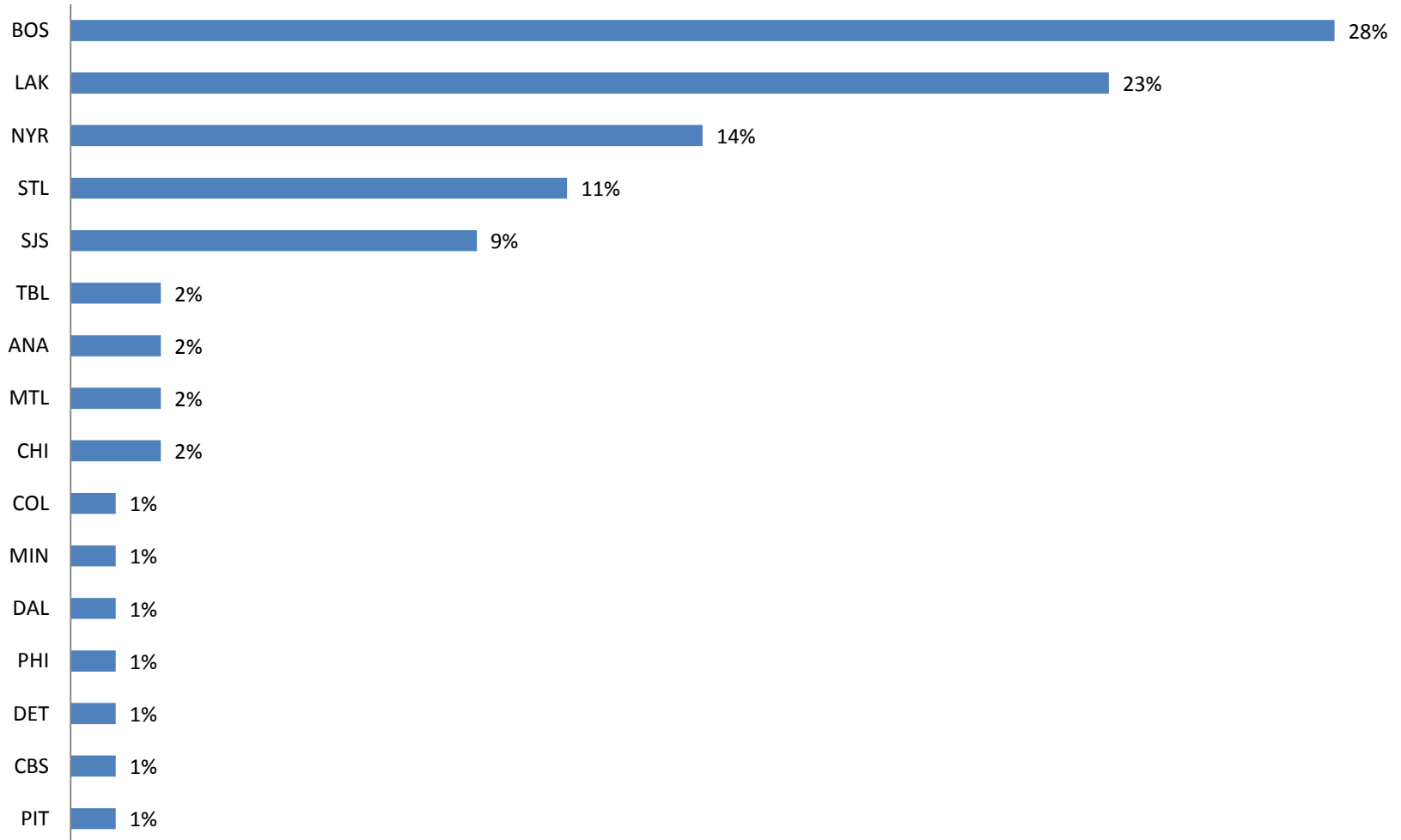## Probability of winning eastern conference

Team	Probability
BOS	44%
NYR	27%
MTL	7%
TBL	5%
CBS	5%
DET	4%
PHI	4%
PIT	4%

Probability of winning western conference

Team	Probability
LAK	37%
STL	24%
SJS	18%
CHI	7%
ANA	6%
MIN	3%
DAL	3%
COL	2%

First round

Stanley Cup probabilities

Team	Probability
BOS	28%
LAK	23%
NYR	14%
STL	11%
SJS	9%
TBL	2%
ANA	2%
MTL	2%
CHI	2%
COL	1%
MIN	1%
DAL	1%
PHI	1%
DET	1%
CBS	1%
PIT	1%

Second Round

BOS 77% — MTL 23%

PIT 26% — NYR 74%

CHI 65% — MIN 35%

LAK 80% — ANA 20%

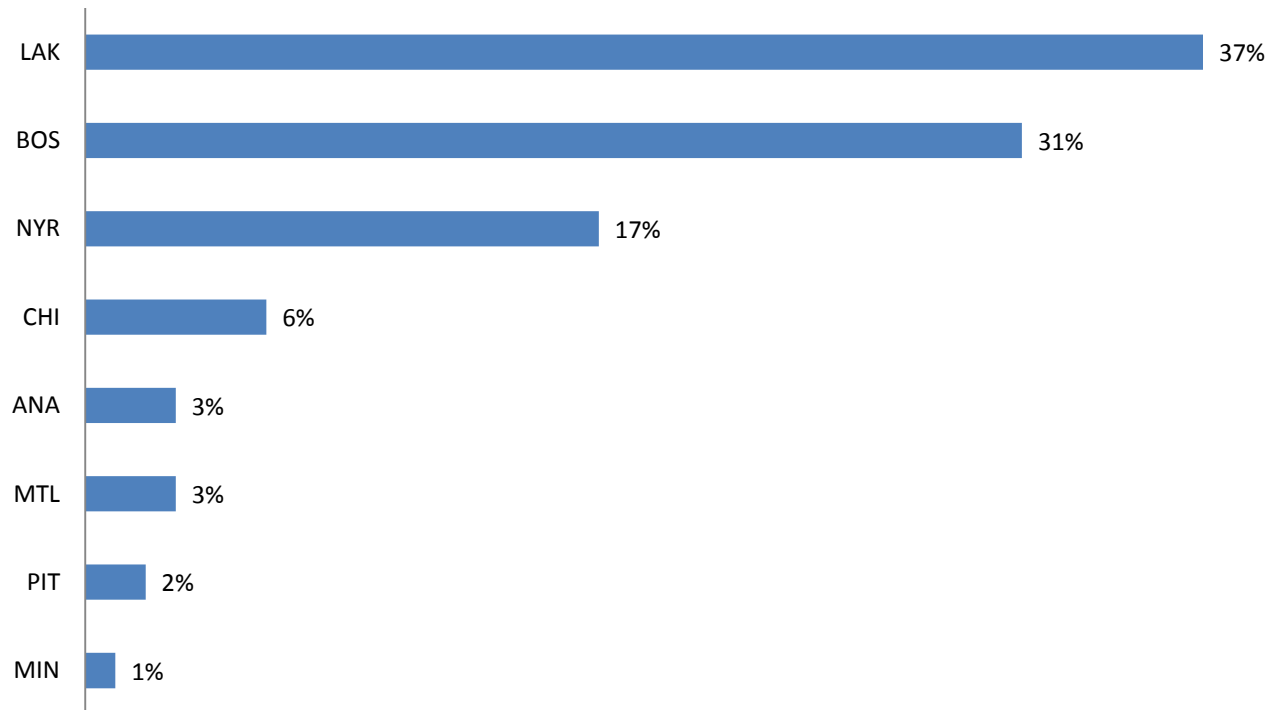# Second round

Probability of winning eastern conference

BOS	51%
NYR	33%
MTL	9%
PIT	7%

Probability of winning western conference

LAK	64%
CHI	19%
ANA	11%
MIN	6%
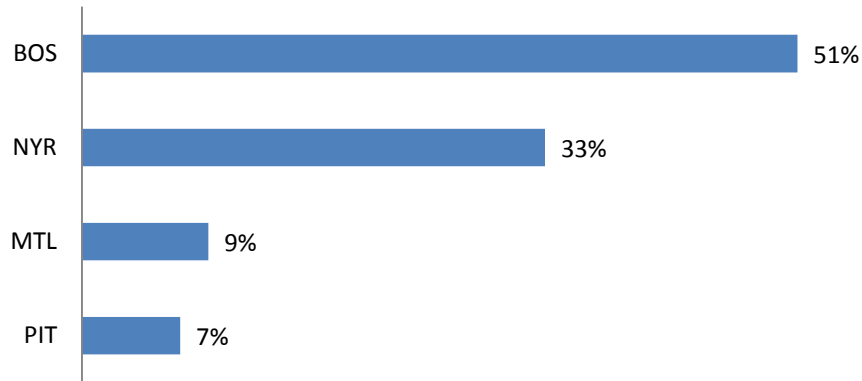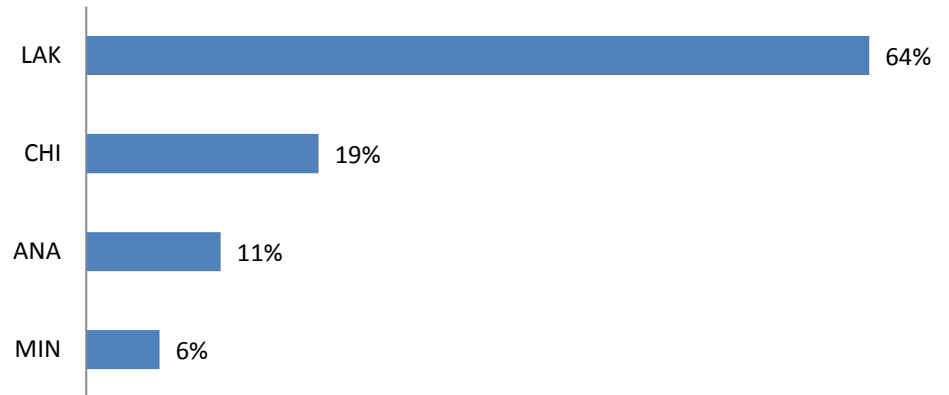
Second round

Stanley Cup probabilities

Team	Probability
LAK	37%
BOS	31%
NYR	17%
CHI	6%
ANA	3%
MTL	3%
PIT	2%
MIN	1%

Conference finals

NYR 68%	**MTL 32%**

CHI 23%	**LAK 77%**

Stanley Cup probabilities team save pct

LAK	51%
NYR	31%
MTL	10%
CHI	8%
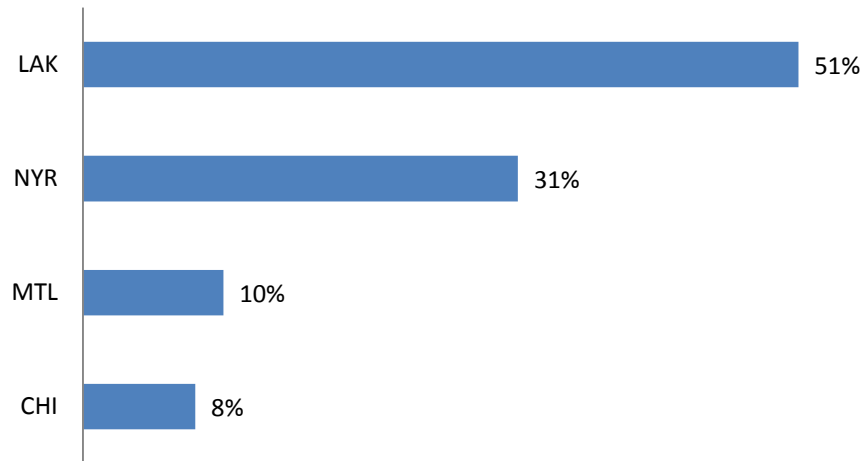
Stanley Cup probabilities goalie save pct

LAK	38%
NYR	31%
CHI	22%
MTL	9%
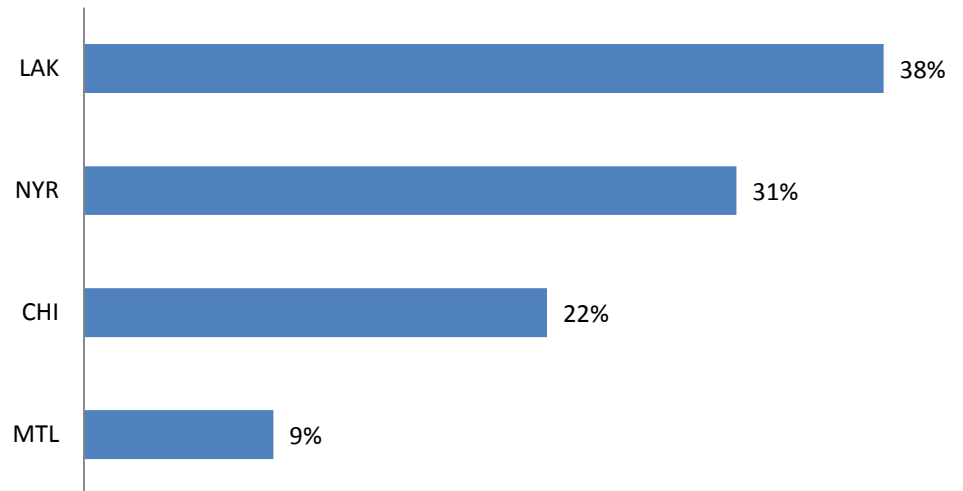
Stanley Cup Final

LAK 61%	NYR 39%

- Limitations
  - Does not account for injuries
  - Does not take into account postseason performance
  - Goalie save percentages
  - Inherent randomness

Next steps

- Identify other variables
  - Zone starts
  - FF/FA separately
  - FF in last 20 games
  - Adjusted save percentage
  - Starting goalie save percentage
  - Possession stats for special teams
- Take postseason results into account