Scheduling Effects in the NBA and NHL

Justin B. Post, Len Stefanski and Jason A. Osborne
Dept. of Statistics, N.C. State University

NC State University
DEPARTMENT OF STATISTICS

2017 New England Symposium on Statistics in Sports at Harvard University, Cambridge, MA

Introduction

PRO BASKETBALL

The N.B.A.'s Back-to-Back Problem: 'Rest? It's Your Job, Man!'

By SCOTT CACCIOLA DEC. 18, 2016

"It's tough because we have a young team," Luke Walton, the Lakers's head coach, said. "You have 19-year-olds who aren't used to this. The mental preparation that it takes and the mental strength that it takes to fight through that fatigue is challenging."

- On average, teams today who played yesterday score significantly less than those that were idle. (Entine and Small (2008).)
- Scheduling discrepancy has rested home teams playing tired visitors. Particularly for NBA West teams. In an attempt to balance this advantage, other teams have reduced *dis*advantage as road teams. (See map courtesy of reddit user: `cs_irl`)
- After adjustment for scheduling discrepancy using factorial effects in linear/generalized linear models, much of the home ice advantage in the NHL vanishes. Some of it remains in the NBA.

Scheduling discrepancy for visitors

- Visitors played yesterday more than **twice as frequently as the home team** (both NHL/NBA).
- NBA teams a little more likely to have played yesterday:

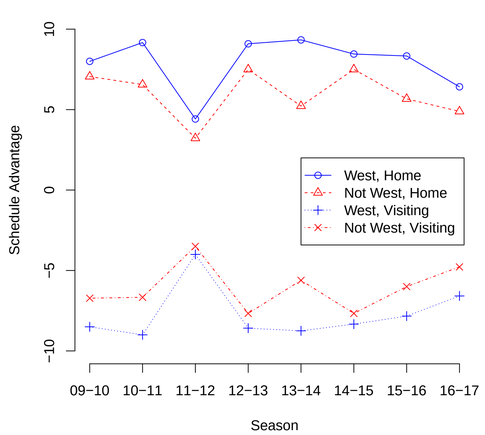
NBA	Visitor		Home	
Home	Played	Idle	Total	
Played	346	332	678	**(14%)**
Idle	1143	3099	4242	
Visitor Total	1489	3431	4920	
	(30%)			

NHL	Visitor		Home	
Home	Played	Idle	Total	
Played	234	280	514	**(10.4%)**
Idle	967	3439	4406	
Visitor Total	1201	3719	4920	
	(24.4%)			

- Effect is most pronounced for GSW,DEN,MIA,UTA,POR,HOU. NBA Regional effects (see map) highly significant ($p < .0001$)

2016-2017 Team	Location	Team rested, opp tired	Team tired, opp rested	Advantage
GS	home	15	0	15
GS	away	3	14	-11
CHI	home	8	3	5
CHI	away	5	10	-5

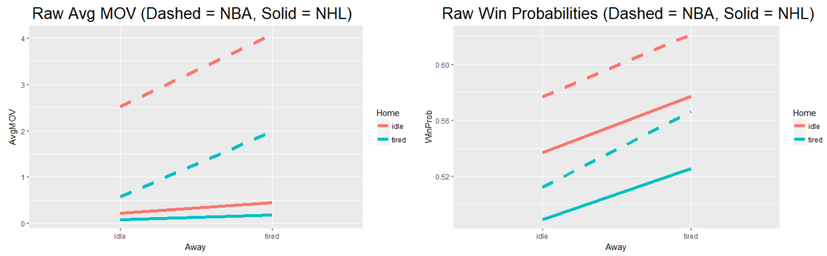
Regional advantages in the NBA



Data Sources

- `pro-basketball-reference.com` - `covers.com` - `hockey-reference.com`

Unadjusted plots:



Linear and generalized linear models

For a given season, linear models for margin of victory Y with factorial effects for yesterday and team:

$$Y_i = \beta_0 + \overbrace{\sum_t \beta_t^H x_{it}^H + \sum_t \beta_t^V x_{it}^V}^{\text{Home, Visiting team effects}} + \underbrace{\beta^{HP} x_{i,31} + \beta^{VP} x_{i,32} + \beta^{HVP} x_{i,33}}_{\text{Fatigue effects}} + \epsilon_i = \mathbf{x}_i\beta + \epsilon_i$$

- Model then pools over last 4 seasons, with team effects nested in season.
- x^H, x^V - indicators for which home, visiting team, $x_{i,31}, x_{i,32}, x_{i,33}$ - indicators for playing yesterday
- Similarly for win probability (π) models, with $\log\left(\frac{\pi_i}{1-\pi_i}\right) = x_i.\beta$

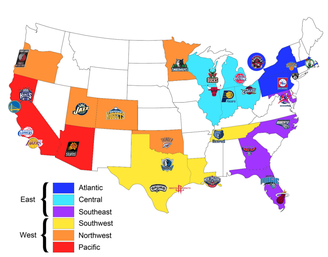
```
proc glimmix data=teamseq; title "yesterday-idle";
   class hteam vteam hidle_yest vidle_yest season;
   *model home_mov=hteam(season) vteam(season) hidle_yest|vidle_yest ;
   model home_wins(event='1')=hteam(season) vteam(season) hidle_yest|vidle_yest /dist=binary;
   estimate "intercept" intercept 1/ilink;
   lsmeans hidle_yest|vidle_yest/ilink;
```

Parameter Estimates

Estimated margin of victory (points/goals):

	Adjusted margin of victory						Adjusted win probability					
	Visitor(NBA)			Visitor(NHL)			Visitor(NBA)			Visitor(NHL)		
Home	Played	Idle	Mean	Played	Idle	Mean	Played	Idle	Mean	Played	Idle	Mean
Played	2.3	0.6	1.5	**0.0**	0.12	0.06	0.60	0.52*	0.57	0.49	0.50	0.49
Idle	3.8*	2.6*	3.2	0.45*	0.23*	0.34*	0.65	0.60	0.63	0.58*	0.54*	0.56*
Mean	3	1.6	2.3	0.22*	0.17*	0.19*	0.63	0.57	0.60	0.54	0.52	0.53*
	NBA(Points)			NHL(Goals)			NBA			NHL		

- After adjustment for team & fatigue effects, estimated NHL home team MOV is 0, win probability estimate is 0.5. Fatigue effects plausibly additive
- Using the fitted models, the best, worst teams, on average:
  2017 Home Court GS Warriors win by 14.8 points, 2017 Road GS Warriors win by 7.6 points
  2014 Home Courts 76ers lose by 10.6 points, 2015 Road 76ers lose by 12.7
  2017 Home Ice Caps win by 1.6 goals, 2017 Road Rangers win by 0.9 goals
  2015 Home Ice Coyotes lose by 1.3 goals, 2015 Road Sabers lose by 1.7 goals



Perceived Schedule Effects

- Fitted SLR of home MOV on Vegas Line (V) in the NBA: $\widehat{MOV} = 0.01 + .95V$
- Fatigue effects on spread highly significant. $r^2 = .82$ of variability in spread explained by linear model. $MSE = 2.9$
- Who's at home explains a bit more variation than who's visiting.

```
The SAS System - The GLM Procedure
Dependent Variable: vegaswinby
                                        Sum of
Source                      DF       Squares    Mean Square  F Value  Pr > F
Model                      302    222536.6167      736.8762    87.03  <.0001
Error                     5816     49241.3050        8.4665
Corrected Total           6118    271777.9216

R-Square     Coeff Var      Root MSE    vegaswinby Mean
0.818818      100.3955      2.909729          2.898268

Source                      DF   Type III SS    Mean Square  F Value  Pr > F
hteam(season)              145   111447.1980      768.6014    90.78  <.0001
vteam(season)              145    98784.6556      681.2735    80.47  <.0001
h_ystrday                    2     1206.0397      603.0199    71.22  <.0001
v_ystrday                    2      602.0291      301.0146    35.55  <.0001
h_ystrday*v_ystrday          4       33.7255        8.4314     1.00  0.4083
```

Overtime effects

- For betting the over/under, including overtime effect debiases estimators of team capacities to score and give up points.
- Forecasting today's game using data til yesterday with/without OT effect

```
%macro betday(season,pickday);
proc glm data=temp; title "overunder - prospective betting";
   iplay_f gdate  <= &pickdate and  season="&season";
   class hteam vteam ; *overtime;
   model gsum2=hteam   vteam overtime2; * gsum2 missing, overtime2=0 today;
   *model gsum2=hteam   vteam;          * both observed yesterday;
   output  out=bets p=p;
%mend;
* loop over season, pickday ;
```

  - Forecasts from models without OT bigger than forecasts from models with OT
  - Use diff between forecast and over/under: $PD5 = round(\widehat{\text{sum}} - \text{over/under}, 5)$, select games where $|PD5|$ large, win with with $\widehat{p} = 54\%$ when using OT effect.

(With overtime)	Counts Without overtime PD5 < 10	Counts Without overtime PD5 ≥ 10	Winning percentage Without overtime PD5 < 10	Winning percentage Without overtime PD5 ≥ 10	
PD5 < 10	3215	174	.503/.505	.454*	
PD5 ≥ 10	256	898	.508/.504	.543*+	.535

+ significantly different from 0.5 ($p = .01$), ∗ winning pct equal for both forecasts

Findings

- Broad takeaways: (1) Fatigue explains some home court and a lot of home ice advantage. (2) Some of the NBA Western conference dominance is attributable to systematic home court scheduling advantage over Eastern conference, with home team a moderately greater determinant of victory than road team.
- Technical takeaways: Schedule effects highly significant on both points/goals and win probability. After adjustment for team strength and schedule effects, estimated home ice advantage for both teams rested estimate is 0. NBA home court advantage remains at $\widehat{\mu} = 2.3$, averaging equally over fatigue conditions. Estimating team scoring/defending effects improved by inclusion of OT effect, preliminary evidence of betting market inefficiency.

References

Entine, O. A. and Small, D. S. (2008), 'The role of rest in the nba home-court advantage', *Journal of Quantitative Analysis in Sports* **4**(2).