Home Sweet Home: Quantifying Home Court Advantages For NCAA Basketball Statistics

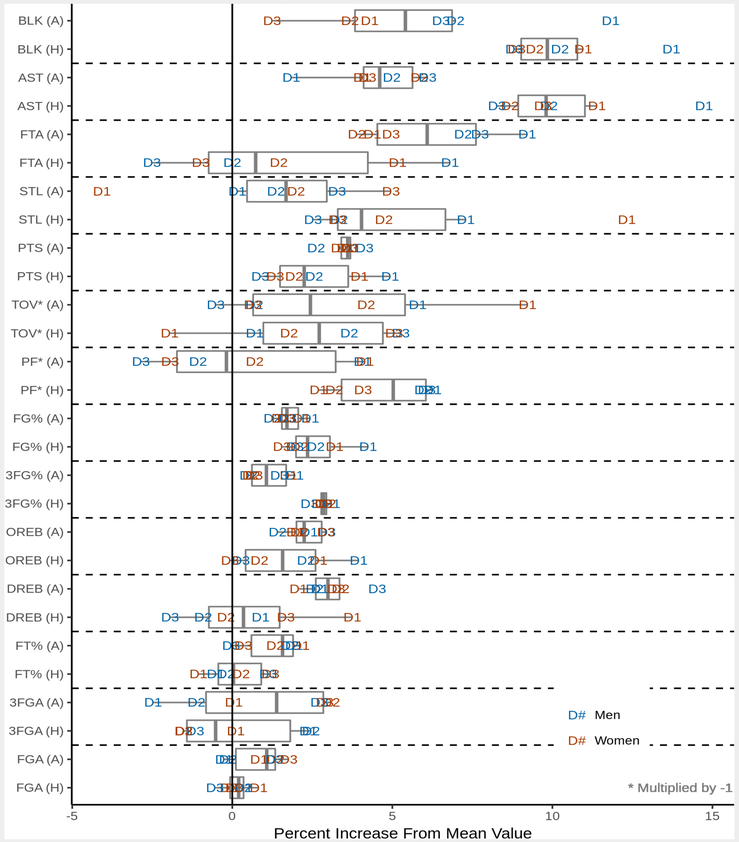
Matthew van Bommel, Luke Bornn, Peter Chow-White, Chuancong Gao / Simon Fraser University

Home vs Away vs Neutral Statistics

Results:

- Home teams outperformed away (and neutral location) teams in nearly all statistics across every gender-division-season combination
- AST and BLK (the most subjective statistics) had the greatest home advantages and much of the advantages remained when compared to neutral location teams
- Home teams experienced a big boost in FG% and 3FG%, while away teams performed slightly worse than neutral location teams
- FT% was negatively impacted for away teams compared to neutral location teams but home teams did not receive an additional boost
- Scorekeepers tended to have greater home team biases when observing men compared to women or higher divisions compared to lower divisions

Using over 100,000 games between the 2011-2012 and 2015-2016 seasons, we compared the difference in home, away, and neutral teams across a variety of box score statistics for both genders and all three divisions. The results are displayed in the following figure.



(A) – neutral teams' advantage over away teams
(H) – home teams' advantage over neutral teams

Attendance Impact on Statistics

Results:

- Referee bias in favor of the home team increased in games with high attendance (statistically significant for all but D2 Women)
- Referee bias had a greater impact on point totals in D2 and D3, compared to D1
- No evidence of attendance impact on FT% was discovered
- No evidence of attendance impact on scorekeepers was discovered

Two-sample t-tests

Mean (Home – Away) Statistic Value in games with **High Attendance** (Top 25%)

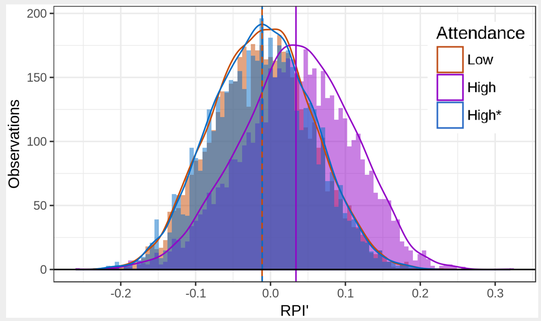
vs

Mean (Home – Away) Statistic Value in games with **Low Attendance** (Bottom 25%)

Controlling for home team strength advantage:

RPI' = Home RPI - Away RPI

RPI = (0.25 * Team Win %) +
(0.50 * Opponents' Win %) +
(0.25 * Opponents' Opponents' Win %)



Statistical matching and bootstrap sampling used to construct new high attendance (High*) samples, with equal RPI' distributions to low attendance.

The matching algorithm has a random component, so we repeated the process 1000 times for each statistic, performing 1000 distinct tests. The average p-value results are displayed in the following table.

		Division I		Division II		Division III	
	Overall	Men	Women	Men	Women	Men	Women
Low	384	1507	427	281	181	181	112
High	1584	7042	1824	897	527	500	300
PF	3.67×10^{-4}	$\mathbf{9.20\times10^{-15}}$	$\mathbf{5.76\times10^{-9}}$	$\mathbf{3.53\times10^{-4}}$	0.002	$\mathbf{4.64\times10^{-7}}$	$\mathbf{1.01\times10^{-12}}$
FTA	0.003	$\mathbf{3.71\times10^{-10}}$	$\mathbf{2.80\times10^{-4}}$	$\mathbf{2.04\times10^{-4}}$	0.024	$\mathbf{2.08\times10^{-4}}$	$\mathbf{1.17\times10^{-6}}$
PTS	0.031	0.056	0.127	$\mathbf{7.79\times10^{-6}}$	0.015	0.001	0.020
3FGA	0.068	-0.014	0.108	-0.025	-0.227	-0.009	0.096
AST	0.071	-0.098	0.130	0.069	0.034	0.006	0.162
TOV	0.072	0.046	-0.086	0.095	0.229	0.026	0.023
BLK	0.092	$\mathbf{2.17\times10^{-4}}$	-0.091	0.001	0.235	0.044	-0.272
FG%	0.111	0.144	0.266	0.002	0.049	0.048	0.269
FGA	0.125	$\mathbf{-1.95\times10^{-4}}$	-0.093	0.080	-0.177	-0.234	-0.289
OREB	0.133	0.312	0.250	0.003	0.035	0.234	-0.094
DREB	0.138	-0.015	-0.150	0.063	0.252	-0.261	-0.227
3FG%	0.158	0.163	-0.041	0.245	0.183	0.188	0.284
FT%	0.164	0.284	-0.252	0.071	0.150	0.241	0.151
STL	0.169	0.169	-0.272	0.183	0.230	0.055	0.277

Low / High – attendance cutoffs
Overall – average p-values across all gender-division combinations
Bold values – significant t-test results at α = 0.0006 (Bonferroni correction)
Negative values – p-values multiplied by -1 when away teams had the advantage

Quantifying Home Court Advantage

Results:

- The home court boosts for AST and BLK were equivalent to having 12 and 43 additional possessions in a game respectively
- Men received an additional home court boost for BLK while women received a lesser home court boost for OREB and FGA
- Team strength advantage coefficients were highly correlated (0.85) with home advantage coefficients, thus scorekeepers may have been more biased in favor of stronger teams

LASSO Poisson regression model

$$\hat{\beta} = \underset{\beta_0,\beta}{argmin} -\frac{1}{N}\sum_{i=1}^{N}\left(y_i\left(\beta_0+\beta x_i\right)-e^{\beta_0+\beta x_i}\right)+\lambda\sum_{j=1}^{p}|\beta_j|$$

Predicting statistic counts while controlling for:

- RPI`
- Possessions
- Gender / Division

After estimating the models, the percentage impact values are computed as e^{β_j}-1 and are presented in the following table.

		BLK	AST	TOV	STL	DREB	OREB	3FGA	FGA
Home	Overall	12.91	12.27	4.48	4.48	1.90	1.74	0.51	0.09
	Men D1	8.87	0.26		-0.34	3.08	0.27	-0.77	-1.00
	Men D2	3.70	1.09	0.89		0.89	-0.57	0.38	
	Men D3	3.55					-0.71	0.59	
	Women D1		0.90		0.22	1.99	1.58	-0.37	0.89
	Women D2	-1.69				-0.03	0.62	0.80	0.82
	Women D3	-3.70	-0.09	0.27	1.56	-0.87	1.83	1.55	1.45
Division	Men D1		2.04	7.02	-7.18	-1.85	-0.86	8.16	-0.89
	Men D2	-8.08	-0.52	4.49	-6.85	-1.39	-1.96	3.81	-0.86
	Men D3	-9.87		0.84	-0.79	1.90	-0.81	0.52	-0.12
	Women D1	5.45	1.27	18.34	19.45	-0.55	16.78	-2.60	1.63
	Women D2	-3.94	-0.81	20.70	19.10	2.46	13.50	-5.90	0.75
	Women D3	1.03	-1.50	30.87	32.16	5.92	22.03	12.36	1.99
	RPI`	10.14	12.27	6.31	10.01	5.61	5.90	0.90	1.71
	Possessions	0.30	0.97	-0.57	1.05	0.46	0.74	2.14	0.80

Home – increase in statistic frequency for home teams compared to away teams
Division – baseline differences in gender-division combinations
Empty Cells – variables not selected in models

To read the full paper visit:
matthewvanbommel.com/nessis
or
arxiv.org/abs/1909.04817

Contact:
matthew.vbommel@gmail.com