

Home Advantage: Does it Exist?

Joel C. Henriques*

May 4, 2020

Abstract

In sports, more than any other profession, there lives anxiety-enhancing hysteria. For the most part, it often takes the form of sayings rather than statistical probability. For instance, in basketball there is the saying commonly known as the “hot-hand” or “heating-up” which infers that a player is most likely to score their next shot after making a few consecutive shots. One of the most promulgated sayings in the realm of sports happens to be the notion of “home advantage”, meaning that a team is playing on their home field and should benefit from it. The purpose of this study is to determine whether home advantage actually exists in the National Basketball Association. I have designed a model that determines the winning percentage through the use of the following regressors: game location, forecast winning percentage, elo skill ranking of a team going into the game, amount of points scored, and whether a game is a playoff or not. It turns-out that the hysteria surrounding the home advantage is false, and the results of this study propose that home advantage certainly increases the odds of winning. One model suggests that the home advantage can lead to a 24.5% increase in the likelihood of winning when only regressing the location of the games over game results. However, models that incorporate more independent variables suggest that the home advantage only increases the likelihood of winning by approximately 1%.

*Department of Economics, University of Oklahoma. E-mail address: Joel.C.Henriques-1@ou.edu

1 Literature Review

Once upon a time, a time before the time of coronavirus, there was this thing we called sports. Individuals would enter into sports arenas to find a sea of faces sweating, crying, and cheering while engaging in an arbitrary activity known as “the wave”. Fans are yelling at sports players, sports players are yelling back at fans. Then, all of a sudden, “Breaking News” alerts flood your iphone about Zion Williamson taking Duke onto the next leg of the NCAA March Madness Tournament. Do you remember these days? Unbeknownst to the sports world was the day that the music would end. In an effort to keep the memories and music going, this study will look into some of the intricacies of what made sports truly special.

Many believed in this idea that sports teams had an inherited advantage through having games on their home field or court. This was commonly known as “home advantage” and this is well known throughout all of sports. A proponent of this phenomenon is Bray (1999) who suggests that in hockey “[o]ver the total 20-year period, a home-game winning average of 52% was present within the league. However, when the record of individuals teams were examined, teams were found to have won 17.3% ($p < .001$) more games at home than away.” In making this statement, Bray (1999) is illustrating that there is a distinct advantage for teams playing on their home fields. This would appear to be logically sound, because teams have much more familiarity with their home fields which could lead to better performances. Also, there is a traveling aspect for the away teams that could make players more fatigued or unable to perform at maximum capacity.

On the contrary, would it be possible to somehow cheat on the home advantage? According to Balmer, Nevill, and Williams (2001), who studied the the impact of home advantage in the Winter Olympics, suggests that “[w]hen events were grouped according to whether they were subjectively

assessed by judges, significantly greater home advantage was observed in the subjectively assessed events ($P = 0.037$). This is a reflection of better performances, suggesting that judges were scoring home competitors disproportionately higher than away competitors.” In other words, Balmer et al.(2001) believed that in sports that required assessments from judges were more likely to be susceptible to cheating by favoring the home countries’ competitors. Thus, the idea of home advantage could be exploited from the fact that referees or judges cheat in favor of the home team. Despite these claims, the Olympics can be viewed as a wildly different scenario than the average NBA game. For instances, Hiller and Wanner (2011) emphasize the political side of the Olympic’s dichotomy when they write that “rather than having an expected universal appeal, the Olympics can be caught in local political debates that reflect criticism and differences in policy priorities, party ideology, and leadership preferences which result in public conflict, opposition, or suspicion about the Games.” Essentially, Hiller and Wanner (2011) are demonstrating that countries that hold the Olympics succumb to governmental scrutiny due to the overwhelming pressure of hosting these games. Therefore, the pressures of cheating could be heightened in the Olympic games due to the political aspects of hosting these games.

Furthermore, if cheating was extremely prominent in the home advantage phenomenon, then those teams that change the locations of their stadiums would not show any signs of adverse effects. According to Pollard (2002), “[h]ome advantage during the first season in a new stadium after the move was significantly less than home advantage in the final season in the old stadium ($P = 0.011$)... It is estimated that about 24% of the advantage of playing at home may be lost when a team relocates to a new facility.” In making this claim, Pollard (2002) suggests that teams that relocate their stadiums will experience a drastic decrease in their winning percentages at home. Thus, there is a minuscule possibility that the home advantage phenomenon can be attributed to

massive amounts of cheating in sports.

2 Description of the Data Used

The data used in this research was all sourced from DataHub. This organization has a robust collection of data across various field such as economic indicators, healthcare, government, sports, and much more. For purposes of this study, the data ranges over the years of 1947 to 2015 of every game played in the NBA (Kariv and Pollock, n.d.). The sample size contains 126,276 observations and 24 variables, however, only 6 variables are of interest for this study (Kariv and Pollock, n.d.). This data used is, indeed, panel data because of the different variables being observed over various time periods.

3 Empirical Methods

I created a few models that regress the game results on the game locations. In addition, the other models take into consideration the utilization of more independent variables such as forecasts, points scored, skill ranking of teams entering into a game, and playoff games; in order to further unravel the relationships that these variables share with winning. The predicted variable, game results, along with the regressor variable, game location, take on binary values. These are dummy variables which commonly obtain a value of one or zero in order to translate qualitative data into quantitative. The following models are performed in logit $[G(x\beta) = \frac{\exp(x\beta)}{1+\exp(x\beta)}]$ and probit $[G(x\beta) = \int_{-\infty}^{x\beta} \Phi(z)dz]$ form.

$$Y_{GameResults} = \beta_0 + \beta_1 game_{location} + \varepsilon \quad (1)$$

$$Y_{GameResults} = \beta_0 + \beta_1 game_{location} + \beta_2 forecasts + \beta_3 points + \beta_4 playoffs + \beta_5 elo + \varepsilon \quad (2)$$

The chart listed below describes each of these variables.

Variables Name	Definitions	Type of Variable
Game Result	Whether a team won or lose	Regressand
Game Location	Identify if the game is home or away	Covariate
Forecast	The likelihood of a team to win	Covariate
pts	Points Scored	Predictor
elo	Teams skill level ranking going into a game	Predictor
isplayoffs	Identify if its a playoff game	Covariate

Table 1: *Variables Table*

The use of these models shall uncover the probability of winning an NBA game due to having the home advantage. Logit as well as the probit models are used for nonlinear models to help approximate likelihood estimates which is an appropriate method for this study. All of the explanatory variables are unit based, except for the independent variable, forecast. This variable is in percentage terms and will be accounted for in the interpretations section.

Furthermore, the primary regressor under evaluation is the game location of each win. The study relies on the relationship between the game results and game locations, because this relationship will give insight into whether home or away games have higher probabilities of winning. In return, it will emphasize if fans are quite delusional or are generally correct about teams benefiting from home advantage.

Currently, the NBA has a schedule of 82 games per team that are separated equally for the number of home and away games. In addition, the playoffs are an extension of the regular season

where half of the teams are chosen to continue to compete for the NBA championship. Many spectators and sports analyst, such as Steven A. Smith, emphatically proclaim that the regular season NBA games barely matter in comparison to the playoffs. In order to observe these claims by “sports savants”, I have integrated the predictor variable playoffs into these models to understand how these may impact winning percentages. If playoffs matter more than the regular season teams, then there should be a positive correlation between the two dummy variables, playoffs and game results.

Another aspect that is worthy of being inspected, is the elo skills ranking going into games. To further explain this variable, Saxon (Nov.14,1992), a prominent writer for the New York Times, writes that “Professor Elo introduced his rating system in the United States [in] about 1950. It has since been adopted not only the World Chess Federation but by sports organizations as well. . . Professor Elo once commented [sic]. ‘It is a means to compare performances, assess relative strength, not a carrot waved before a rabbit.’” Essentially, Saxon (Nov.14,1992) is illustrating that this is a measurement used to quantify skills in the sports world, however, the elo measurement is still susceptible to those who maximize the value of their skills through “stat-packing”. For instance, people claimed that Russell Westbrook was “stat-packing” whenever he went for an uncontested rebound, because the value of his skills would inherently increase from the numbers of times, he received a triple-double per game. In any event, all measurements are merely imperfect but still quite interesting to examine. If teams are awarded off their ranking going into a game, then they might be prone to the idea of maintaining the reputation. This was seen by the 73-win-team Warriors who had appeared to be chasing the best NBA win-record than a championship (since they lost to the LeBron . . . okay technically the Cleveland Cavaliers, but we all know it was mainly to LeBron). Therefore, teams might be psychologically impacted by these measurements, since

these metrics make for excellent headlines in the media and could lead to pressure from general management who seek to maximize profits from the publicity.

In fact, the independent variable forecast will also be following the same logic as the elo skills measurement. Many analysts predict or forecast the likelihood of teams winning verse other teams in the league. However, there was a massive scandal in the NBA, where the refs were getting paid to throw games for those who were betting on certain teams to win. A reporter for ESPN, Eden (Feb.19, 2019) explains that “bribe[s] [were] only two dimes, \$2,000 per game – an outrageous bargain. If the pick won, the ref got his two dimes. If the pick missed, the ref owed nothing; Battista would eat the loss. A ‘free roll,’ as they call it. But this referee didn’t lose much. His picks were winning at an 88 percent clip, totally unheard of in sports betting.” In other words, Eden (Feb.19, 2019) is expressing the depths of the corruption happening in the NBA through referee’s ability to manipulate the game via the whistle. This way, the referees have an incentive to call fouls in favor of the gamblers, since they have a stake in the profits. Thus, evaluating the link between the forecasts and the winning percentage might shine some light on the scandals that have taken place within the NBA. Finally, the last independent variable is points scored. The purpose of this variable is for good measure, because it would be logical to think that the more points scored would lead to higher winning percentages.

4 Research Findings

This study utilizes the methods logit and probit in order to assess the affects of the home advantage phenomenon on winning percentage. The majority of the interpretations will come from Table 2-3 which are marginal effects. The importance of looking at the marginal effects is that it will

allow the results to be reported in unison of their impact on winning percentage, even though each independent variable is in different units. Meanwhile, Figures 1-2 are for those who are interested in looking at the fixed effects that better corrected for serial correlation and heteroskedasticity.

In Panel A of Table 2, the results suggest that for every NBA game played on a team's home court, will have approximately a 1%(0.009) increase in the probability of winning the game. This parallels the results of the relationship between points and game results. For each point that is scored, will increase the likelihood of winning by around 1% (0.009) as well. This makes sense since there are around 80-100 points scored per team in an NBA game. For instance, if the Chicago Bull's scored 90 points that equates to a 81% ($(0.009 \times 90 = 81\%$ but rounded $= (0.01 \times 90 = 90\%)$) change of winning the game. In order to try to grasp the effects of gambling and cheating in the NBA, I incorporated the forecasts predictions. The results suggest that the probability of winning raises by 0.007% for every additional percentage point increase in the forecast predictions. The Elo rating is a ranking of a team's skills going into a game. The Elo ranking has virtually zero ($8.264e-05$) impact on the probability of winning. Even though the Lakers may have had their best 5 games in a row and their skill level is considered extremely high, these past performances would not have much of a bearing on the likelihood of winning their next game. The Elo skills follow the same fallacy trends as the "Hot Hand". Lastly, the binary regressor variable Playoff game is used to see if these games truly matter much more than regular season games. The answer is ... yes of course these games matter more, and the data also supports this by suggesting that each game that is identified as a playoff game will increase the chances of winning by approximately 1.5%. All of these relationship were statically significance at the 1% level. As for Table 3 is also marginal effects but only for regressing game location over game results. Taking into account only game location; each home game played, will increase the likelihood of winning by 24.5%.

5 Conclusion

Ultimately, sports fans are not as delusional as they may appear (Except for Buffalo Bill's fans). This study proposes that the home advantage phenomenon truly exists in the NBA and will raise a team's probability of winning by around 1%. As far as cheating having a meaningful impact on the NBA, it is highly unlikely. Even if the team had a predicted chance of 80% to win a particular game, this would only equate to 0.006% ($0.00763 \times .80 = 0.0061$) chance of actually winning. The largest takeaways from this study, in addition to the existence of home advantage, is the 1 for 1 ratio between points scored and winning percentage as well as playoff games truly mattering more than NBA regular season games. The importance of this is seen in the situation where a team is playing on their home court in a playoff match and has scored 112 points. This would mean that this team would have a 103% ($((112 \times 0.009) + (1 \times 0.009) + (1 \times 0.015)) = 1.032$) chance of winning the game. However, this data set of the 126,276 observations over the years spanning from 1947-2015, shows that the probability of scoring over 112 points in a game is about 49% ($62442/126276 = 0.494$). Hence, the fact that teams could have all the variables in their favor, but still manage to fall sort of securing a 100% victory. This is why sports are truly spectacular, because anything can happen.

6 Recommendations For the Future

This research will suffer from the omitted variable bias due to nature of the study, since there is no way to possibly account for everything that contributes to winning in sports. However, the fixed effects and marginal effects should have limited the effects of this error. More studies in the future should look deeper into variables that may impact cheating in the NBA other than forecasts. I think forecasting was simply an interesting variable to include in this study, but it was not the

primary focus. Also, other models should be used such as machine learning algorithms to possibly determine the ranges of forecast which could be identified as cheating zones.

References

- Balmer, Nigel J., Alan M. Nevill, and A. Mark Williams. 2001. "Home advantage in the Winter Olympics (1908-1998)." *Journal of Sports Sciences* 19 (2):129–139. URL <https://doi.org/10.1080/026404101300036334>. PMID: 11217011.
- Bray, Steven R. 1999. "The home advantage from an individual team perspective." *Journal of Applied Sport Psychology* 11 (1):116–125. URL <https://doi.org/10.1080/10413209908402954>.
- Eden, Scott. Feb.19, 2019. "How former ref Tim Donaghy conspired to fix NBA games." *ESPN* URL https://www.espn.com/nba/story/_/id/25980368/how-former-ref-tim-donaghy-conspired-fix-nba-games. Accessed April 22, 2020.
- Hiller, Harry H. and Richard A. Wanner. 2011. "Public Opinion in Host Olympic Cities: The Case of the 2010 Vancouver Winter Games." *Sociology* 45 (5):883–899. URL <http://www.jstor.org/stable/42857463>.
- Kariv, Adam and Rufus Pollock. n.d. "Nba Elo." *DataHub* .
- Pollard, Richard. 2002. "Evidence of a reduced home advantage when a team moves to a new stadium." *Journal of Sports Sciences* 20 (12):969–973. URL <https://doi.org/10.1080/026404102321011724>. PMID: 12477006.
- Saxon, Wolfgang. Nov.14,1992. "Prof. Arpad E. Elo Is Dead at 89; Inventor of Chess Ratings System." *The New York Times* URL <https://www.nytimes.com/1992/11/14/obituaries/prof-arpad-e-elo-is-dead-at-89-inventor-of-chess-ratings-system.html>. Accessed April 22, 2020.

Figures and Tables

Table 2: Estimations of the Logit Model and Probit with Multiple Independent Variables

Panel A: Logit and Probit Marginal Effects

	Est.	Std. Dev.	z	P> z
Game locationh	0.009	0.003	3.067	< 0.002 ***
Forecast	0.763	0.010	71.49	< 2.2e-16 ***
Points	0.009	0.000	83.81	< 2.2e-16 ***
elo rating	8.264e-05	0.000	5.386	< 7.223e-08 ***
Playoff Games	0.015	0.005	2.893	< 0.004 ***

Notes: Data ranges over the years 1947-2015. Sample size for all variables in Panel is $N = 126,276$.

Table 3: Estimations of the Logit and Probit Model Only with Game Locations

Panel B: Logit and Probit Marginal Effects

	dF/dx.	Std. Dev.	z	P> z
Game locationh	0.2451	0.0027	89.818	< 2.2e-16 ***

Notes: Data ranges over the years 1947-2015. Sample size for all variables in Panel is $N = 126,276$.

Dependent variable:	
wins	
home	0.286 (4.615)
forecast	-1.053 (18.227)
pts	7.366*** (0.000)
elo_i	0.003 (0.045)
is_playoffs	0.000 (0.000)
Observations	126,276
R2	0.500
Max. Possible R2	0.500
Log Likelihood	-1.891
Wald Test	0.010 (df = 5)
LR Test	87,524.070*** (df = 5)
Score (Logrank) Test	41,057.050*** (df = 5)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 1: Fixed Effects with Multiple Variables

Dependent variable:	
wins	
home	0.500*** (0.008)
Observations	126,276
R2	0.030
Max. Possible R2	0.500
Log Likelihood	-41,848.770
Wald Test	3,713.320*** (df = 1)
LR Test	3,830.309*** (df = 1)
Score (Logrank) Test	3,791.422*** (df = 1)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 2: Fixed Effects with Only with Game Locations