

The Home-Court Advantage Effect in NBA Basketball

by Josh Carr

4210008

Major Paper presented to the
Department of Economics of the University of Ottawa
in partial fulfillment of the requirements of the M.A. Degree
Supervisor: Professor Marc Lavoie

ECO 6999

Ottawa, Ontario
April 2015

1. Introduction

Modern sports analytics first rose to prominence in Major League Baseball, based primarily off the statistical innovations of pioneering author Bill James. In 1977, James self-published what would become an annual book series, *The Bill James Baseball Abstract*, where he introduced some radically new statistics to evaluate player and team production. James named his approach 'sabermetrics' in reference to the Society of American Baseball Research. Sabermetrics helped transform baseball analysis into an accepted science (Puerzer, 2002). His concepts were popularized in the 1990's when Oakland Athletics general manager Billy Beane based his team building strategy primarily off sabermetrics. Beane's success, and the initial resistance he faced, is chronicled in Michael Lewis' book *Moneyball* and later in a Hollywood movie of the same name. His strategies are now considered a key driver in the evolution of advanced metrics.

Baseball has attracted analytics research in past decades because data can be collected incrementally. The sport is played one pitch at a time, and with each pitch comes a result that can be viewed as an isolated incident. One can record a large amount of data based on every pitch during the game. The name of the pitcher, the type, location, and result of the pitch can all be recorded and are fairly obvious to the trained eye. Baseball, unlike the other three major North American sports, is not considered a free-flowing game. The discrepancy of the speed and nature of the sports meant that football, hockey and basketball have lagged behind in analytics, instead utilizing only statistics that are observable and easily measured.

Professional hockey is developing its own analytics movement based on the fundamental question James asked decades ago. For years, player contributions were graded by results-based statistics like goals, assists, and saves. In the past, the game simply moved too fast for statisticians to record all the happenings on the ice. More recently, data-collection technology and computing power are becoming

available to allow teams to explore more advanced concepts. Teams are beginning to employ professional statisticians to decipher methods of measuring concepts like player contribution in terms of efficiency or scoring chances generated, rather than simply measuring post hoc data.

The National Basketball Association has also followed Major League Baseball's in transforming data analysis. New-age analytics are beginning to gain traction in NBA front offices and sports media with innovative methods capable of providing a competitive advantage. In today's NBA, it is not uncommon for organizations to employ strategies based on strictly scientific, numbers-based methods.

The beginning of the basketball analytics movement can probably be traced back to Dean Oliver's 2004 groundbreaking book *Basketball on Paper: Rules and Tools for Performance Analysis*. Today, measures such as 'player efficiency rating' and 'value above replacement player' are readily available and far more explanatory than the raw numbers that used to pop up in your father's morning newspaper. This paper will attempt to capture the effect of home court advantage, but where it deviates from previous works is its inclusion of a prominent player evaluation statistic - wins shares. The "holy grail" of a single player evaluation rating is elusive, but those who study the game are employing modern data processing methods to inch closer to capturing it (Oliver, 2004). Win Shares is not a perfect player metric, but it captures a player's contribution to victories in a single, easy to digest statistic.

It is then with a slight sense of irony that modern evaluation methods can be used to address a question that is nearly as old as the game itself. The concept of home advantage is universal across all major North American sports. Although the home advantage has trended downward since the league's inauguration, home teams still tend to win roughly 60% of NBA games (Pollard and Pollard 2005). During the 2013-2014 NBA season, not one team performed better away than at home. In fact, you would have to go back to the 2011-2012 season to find only one team with a better record on the opponents' floor than their own.

The NBA provides a simple opportunity to estimate the home advantage effect, relative to the other 3 North American major sports, since the rules do not change for the home team. In the NHL, home teams are granted an information advantage both when deploying lineup changes - the “last change” rule - and in face-offs. Home team coaches are permitted to make a decision on line changes after the visiting team has deployed their personnel. It is assumed this information discrepancy gives the home coach an advantage. Disentangling the value of the rule with the value of a coach’s learned skill in exploiting the rule is likely very difficult. A similar information gap exists when players are preparing for a face-off. The home team player is permitted to place his stick after the visiting player prepares for the draw. Doyle and Leard (2011) found that the home team wins 2.4 more face-offs per game, likely as a result of the advantage the rule provides.

The MLB dictates that the home team bats last in every inning and only bats in the ninth, or extra innings, when losing or tied. When knowing the order the teams will hit, managers are capable of making offensive and defensive adjustments based on their home/away status. Like the NHL, a study in baseball is likely to require a method to control for a manager’s aptitude in making lineup decisions to take advantage of the rules. Also, the MLB does not hold a standard for stadium dimensions and teams may tailor their personnel to exploit any advantage their home stadium may provide. Any attempt to estimate the home advantage in baseball must effectively control for these factors.

An 82-game NBA season provides a sufficient sample size - unlike the 16-game NFL schedule - to address the question: can the home-court effect be measured?

It is unanimously accepted that playing on one’s home-court provides an advantage. Players and coaches speak to its importance. Teams compete for it in the playoffs. Gamblers handicap decisions on it. What is not fully understood is the value of this competitive edge. Presumably factors such as familiarity with the environment and reinforcement from a friendly crowd - facets of the game that are not directly measured - have positive implications. Other underlying factors such as territoriality, culture, history, and playing conditions are either ambiguous or unlikely to have significant influence in a modern, indoor

sport such as NBA basketball (Swartz and Arce, 2014). Attempts to deconstruct to the phenomena are likely to be complex and difficult to control for. This paper is not so ambitious to determine *why* the home team experiences an advantage but specifically *what* that advantage is. The study will attempt to capture this effect by teasing out the values of other variables affecting the outcome of a game. Isolating the value of home-court advantage will contribute to the increasingly complex decision-making processes of NBA organizations.

This paper will explore some of the contributing factors to the experienced home advantage across all sports. Additionally, it will review previous attempts to quantify these factors both through observed outcomes and econometric means. The literature on quantifying the home effect specifically in professional basketball is relatively scarce and as such, this paper draws upon material in works across several sports when forming its own methods.

This study will address questions around the home advantage in NBA basketball through a series of OLS regressions. I intend to quantify the effect and then test the notion that the absence of travel - rather than the location of the game - is a contributing factor in the advantage. Further, I will explore the distinction between regular season and playoff games when evaluating the home effect.

2. Literature Review

Studies have postulated a variety of factors influencing enhanced performance at home. Courneya and Carron (1992) consolidated the components of the home advantage effect in their 1992 sports psychology literature review. They categorize five major components to the home advantage: game location, game location factors, critical psychological states, critical behavioural states and performance outcomes.

Game location simply refers to the physical location of the game. Games played on neutral sites are not expected to present an advantage, even though one team is designated as the home team.

Game location factors represent crowd, learning, travel, and rule factors. Crowd factors are intended to reflect the positive reinforcement received from a supported crowd. Learning factors include the familiarity of the home team's playing facility. Travel factors account for the implied mental and physical fatigue and disruption of game-day routines from travel. Rule factors apply to some sports in which rules change for the home team - most noticeably batting last in baseball.

The game location factors are considered to influence *critical psychological states* and *critical behavioural states* of the three primary actors involved in the outcome - coaches, players, and officials. Coaches and players historically report that they feel more confident and aggressive when playing at home.

Performance factors are then the measured outcomes that are the potentially influenced components of home advantage.

Courneya and Carron presented the following winning percentages for home teams: baseball - 53.5%, American football - 57.3%, hockey - 61.1%, basketball - 64.4%, European football - 69%.

Nevill and Holder (1999) used Courneya and Carron as a guide to provide a convenient overview of studies on home advantage across all sports. Their paper does not directly address a home estimate in basketball, but rather summarizes how various factors affect home advantage. Greer (1983) assessed that crowd behaviour in college basketball, specifically booing the away team, has a significant impact on home performance. Pollard (1986) found that distance travelled for the visiting team was not significant in professional European football. Jurkovic (1983) found that a majority of college basketball players reported that they feel better and more confident when playing in front of an active and supportive crowd. Pollard (2002) found a decrease of 2.67% when studying home winning percentage the season after a team moves to a new stadium in North American hockey, basketball, and baseball. It is clear that exploring the question of why the home team has an advantage has many layers and attempting the

answer it in one study is not feasible. These findings are helpful in providing context to this study, however, most were not utilized explicitly to pursue a point-estimate of the home effect.

With a basic understanding of the factors that contribute to the home advantage, it is now worth exploring some previous efforts in quantifying the effect. Doyle and Leard (2011) attempted to capture the home effect (along with effects of fighting and momentum) in the NHL by developing a probit model using data from the 2007-2008 season. In their model, the dependent variable was a dummy where 1 represented a win for the observed team and 0 indicates a loss. Their key independent variable, home, was also a dummy variable taking on the value of 1 when the observed team played at home. The remainder of their independent variables included several team level statistics reflecting offensive and defensive performance. The results showed that the home team has a probability of winning the game 54.1% of the time. They acknowledged that this figure is lower than the Courneya and Carron figure of 61.1%, but argued that figure was obtained by simply observing results over several seasons in the 1970s and 1980s.

This estimate of home advantage faces some problems when the model controlled for face-offs won. In the NHL, the home team's centre has the advantage of placing his stick after his opponent prior to a face-off. With this information advantage, the home team player wins 2.4 more face-offs per game and by extension, gives his team more opportunities with the puck. When the model controlled for face-offs won, it was found that the home effect reduced greatly and became statistically insignificant. The Doyle and Leard study did not control for the rule that gives the home team the strategic advantage of changing lineups after the visiting team. This discrepancy in information gives the home team's coach the opportunity to react to opponent's personnel decisions when deploying lineups. It is assumed that this rule, in addition to the face-off rule, provides the home team with a tangible rule-based advantage that can be very difficult to control for.

The change in rules for home teams is most apparent in Major League Baseball. In the MLB, the home team bats in the bottom of the inning and will only bat in the ninth inning when tied or trailing in a

game. These conditions provide the home team with an advantage when employing offensive strategies at the end of games (Shmanske and Lowenthal, 2009). Another consideration is that MLB stadiums are not uniform in their dimensions. Teams have the opportunity to construct lineups to maximize the potential of their home playing environment. For example, Yankee Stadium is famously a “hitter’s park” because of its relatively small dimensions and low outfield wall, so the New York Yankees are wise to spend exorbitant amounts of money on the league’s best power hitters.

Shmanske and Lowenthal (2009) attempted to capture the strategic advantage a home team experiences when batting last in the last inning of a game. They worked off the premise that while the home team can optimize its batting options, the away team holds an advantage on the defensive side by deploying the best fielding and pitching options for the situation. The study attempts to capture the discrepancy in these two strategic factors through a series of logit models. When they isolated potential last innings of a tied game, they found that the visiting team was able to prevent runs at a 2.1% rate greater than the season average, indicating that the away team’s defensive adjustments outweigh the home advantage. The authors suggest that these defensive adjustments may explain why home advantage in the MLB is low relative to other major sports. The home effect exists; however, it is very difficult to disentangle the strategic effects and the psychological effects of game location.

Wang et al. (2011) conducted a study on American College Football to determine the home advantage effect. Their framework is one that is similar to the one practiced in this paper. The authors utilize point differential as the dependent variable as game outcome. Some explanatory variables include game location, season winning percentage, attendance, penalty differential, and distance travelled. To find the effect of home advantage, they took a multi-level regression approach by introducing independent variables in progressive stages. In their most complete model, Wang et al. concluded that the home team has a 5.94 point advantage. The multilevel approach used will be replicated in this study.

When exploring the effect in basketball, it is little surprise that Courneya and Carron included crowd factor as a primary influence on the home team. It is nearly considered a fan's duty to their team that they support as loudly as possible when in the arena. Strong fan bases take great pride in building a stadium atmosphere that is imposing and hostile to the visiting team. In post-game interviews, players and coaches often credit enthusiastic fans for motivation during the game. Players, however, are not the only actors influenced by fan reinforcement. Referees, despite their best efforts, are subject to fan influence. Prince et. al (2012) found evidence that ref bias exists and favours the home team. They found that the home team experiences an 11% advantage in what they call "discretionary turnovers", those of which are heavily subjected to referee judgment (e.g. traveling violations and offensive fouls) and 3% advantage in "non-discretionary turnovers," that are less influenced by referee judgment (e.g. possession after out of bounds). The advantage in discretionary turnovers increases by nearly 10% for the home team during the playoffs, where it is assumed the raised stakes and intensified crowd reactions amplify the pressure on the referee.

Entine and Small (2008) presented an estimate for home-court advantage in basketball when exploring the subject. Their work tests the role of rest on the home advantage and attempts to capture the value of discrepancies in the number of days of rest between home and away teams. They present an estimate for home-court advantage as 3.24 points per game. While their point-estimators for rest are not particularly reliable, their paper served to inspire the basic statistical framework of this study.

There are relatively few studies aimed at estimating the home-court advantage in NBA basketball through econometric means. This paper intends to obtain a point-estimator for the home effect, as well as the effects of days rest, team strength and distance travelled. Where I intend to add most significantly to the literature is with this paper's inclusion of an advanced player evaluation metric - win shares. If the estimates for win shares on game outcomes prove robust, the study may provide evidence of the reliability of the player evaluation measure on team performance.

3. Empirical Strategy

To address the questions surrounding the home advantage, I relied primarily on the open-source statistics archive website, www.basketball-reference.com. The site provides comprehensive data on professional basketball games, players, and seasons dating back to 1946. The site's founder, Justin Kubatko, is a former statistics lecturer at Ohio State University and statistics consultant for the Portland Trail Blazers and currently works as a software consultant for the NBA. Kubatko's website provides a source for both NBA fans and researchers to study the game.

I examined all NBA games during the 2013-2014 season. The league's 30 teams each played an 82-game season, for a total of 1230 games (the league managed to make up the Minnesota-San Antonio game in Mexico City that was infamously cancelled due to unfit stadium conditions). Of course, each of these games includes a "home" and "away" team and an outcome that designates a winner and a loser. Because a game is a competitive contest between two teams, my approach is to estimate team variables on relative terms whenever possible. Constructing the independent variables in the dataset in this way is simple enough, however an issue arises when you consider the dependent variable: the outcome of the game.

It is clear that one game's outcome tells two very different stories depending on team perspective. By simply observing each game from each team's perspective, the dataset will consist of 2460 observations with the home and visitor entries effectively being mirror images of one another with the sum and mean of some independent variables equal to zero. To correct this issue, I have followed the Doyle and Leard (2011) method of selecting an "observed" team and "opponent" for each game, with the observed team being the home team in 615 observations, that is, in half of the games. When assigning the observed team, I ensured each team is represented 41 times in the dataset - at least 20 home games and 20 away games. These 41 observations per team are comprised of 26 games versus in-conference opponents

and 15 out-of-conference opponents in order to hold true to the proportion of games over a full season. In assigning the observed team, I did so without knowledge of the outcome of the game or name of the opponent.

To estimate the value of the home effect, and further investigate its implications, I present the findings through a series of OLS regressions. The regressions are presented as follows:

- 1) What I call the “base model” is presented in progressive stages where one explanatory variable is added at a time in order to illustrate evidence of covariance. Home is presented as dummy variable.
- 2) Teams are divided into high and low travel groups in separate regressions.
- 3) The home team is redefined as a team with zero travel prior to the game.
- 4) The home variable is changed to reflect the length of consecutive home/away games played.
- 5) The base model is repeated for playoff games exclusively.

The base model was established with 5 independent variables and a dependent variable. I will outline each below:

Outcome, as the dependent variable, is measured as the point differential of the game. By adopting the point differential method, the model is measuring the *degree* of victory rather than a simple win/loss statistic. Following the method of Entine and Small (2008), expressing the outcome in this way is expected to yield results for the home effect in points-per-game rather than a change in probability of victory.

Team Strength (TS) is an indicator of the of team success. TS in game n is defined as:

$$(\text{sum of points scored in games 1 to } n-1 / n-1) - (\text{sum of points allowed in games 1 to } n-1 / n-1)$$

TS is a macro approach to capture a team's quality by observing the outcome of their games. There is a large number of team statistics that can be used to explain how a team achieved results, but for the purpose of this study, I am not interested in why a team succeeds or fails on a game to game basis. Rather, I only require a simple comparative measure of quality between the two teams.

Home is = 1 when the observed team plays in their home stadium and = 0 otherwise. It is worth noting that the league's two Los Angeles based teams are an anomaly since they share the same arena. In games between the Lakers and Clippers, the home variable reflects the designated home team per the NBA schedule.

Days Rest (DR) is the number of days between games. It does not apply to the first game of the season.

All games immediately following the NBA All-Star break in February reflect an inordinate amount of rest and will be excluded from the regressions.

Win Shares (WS) are an advanced player metric to capture an individual's overall value to his team.

Players tend to impact team success in a variety of ways based on position, size, skills, offensive and defensive contributions. For this study, I am making the assumption that all win shares are created equal.

Team WS in a particular game is the sum of player WS for all those available to participate in that game.

The WS of injured, suspended, or otherwise inactive players are not included in that game. The statistic was first developed in baseball by Bill James (2002) and later adopted in basketball based on the work of Dean Oliver (2004). Calculating the stat is fairly complex and described in further detail in Appendix 1.

Distance (D) is the distance in kilometers the team travelled to get to the game. Distance can be used as a proxy for travel time and assumed to capture the magnitude of discomfort or fatigue a team may feel when playing away from home. It is notable that the home team does not necessarily show $D = 0$, as the home team may be required to travel home for games following an away game. $D = 0$ only in games where a team plays two or more consecutive games at home.

4. Empirical Results

To investigate the explanatory power of the model and independence of the variables, the OLS regressions were ran in progressive stages. The observed team is denoted ‘t’ and the opponent denoted ‘o’. Variables were introduced in the order they are assumed to have greatest value on outcomes. The sequencing is found below and the results of the first four regressions are shown in Table 1:

- (1) $Pts_t - Pts_o = C + b_1(TS_t - TS_o)$
- (2) $Pts_t - Pts_o = C + b_1(TS_t - TS_o) + b_2Home_t$
- (3) $Pts_t - Pts_o = C + b_1(TS_t - TS_o) + b_2Home_t + b_3(DR_t - DR_o)$
- (4) $Pts_t - Pts_o = C + b_1(TS_t - TS_o) + b_2Home_t + b_3(DR_t - DR_o) + b_4(WS_t - WS_o)$
- (5) $Pts_t - Pts_o = C + b_1(TS_t - TS_o) + b_2Home_t + b_3(DR_t - DR_o) + b_4(WS_t - WS_o) + b_5(D_t - D_o)$

Table 1: Four simpler models of home advantage

Equation	Variable	Coeff	t
1	Constant	0.540	1.491
	$TS_t - TS_o$	0.710	14.25

Equation	Variable	Coeff	t
2	Constant	2.215	-4.44
	$TS_t - TS_o$	0.718	14.769
	Home	5.52	7.816
3	Constant	-2.031	-4
	$TS_t - TS_o$	0.720	14.823
	Home	5.149	7.031
	$DR_t - DR_o$	0.617	1.896
4	Constant	-2.017	-4.046
	$TS_t - TS_o$	0.599	11.816
	Home	5.233	7.279
	$DR_t - DR_o$	0.569	1.783
	$WS_t - WS_o$	0.073	6.915

(1) $R^2 = 0.145$, $F = 203$

(2) $R^2 = 0.187$, $F = 137$

(3) $R^2 = 0.189$, $F = 93$

(4) $R^2 = 0.221$, $F = 84$

The base model, given by regression (5), then reads as shown in Table 2:

Table 2: The base model of home advantage

Variable	Mean	SD	Coeff.	t
Outcome	0.464	13.535		
Constant			-2.03	-4.01
TS _t - TS _o	-0.108	7.271	0.600	11.816
Home	0.500	0.500	5.390	6.898
DR _t - DR _o	0.003	1.126	0.555	1.731
WS _t - WS _o	-0.095	34.846	0.073	6.924
D _t - D _o	23.432	1247.466	1.6x10 ⁻⁵	0.516
n = 1195*	R ² = 0.221	F = 67.435	nHome = 597	nAway = 598

* The first game of the season and the first game following the All-Star break, for either team, are excluded.

The results of each regression yield positive coefficients where assumed. Each of the independent variables, with the exception of distance, is expected to be positive contributors to team success. It is seen that the effect of distance travelled is negligible and not significant. These findings are not consistent with assumptions and an alternative approach will be taken in the next section.

Variables were added sequentially in an effort to identify evidence of collinearity. It is assumed that the number of days of rest and the home variable are completely independent as exogenous variables that are imposed on teams by the NBA's schedule. These assumptions are supported by the relative stability in these variables in the regressions. The variable of Team Strength, however, undergoes a 17% change (from 0.72 to 0.60) with the introduction of Win Shares. The evidence for collinearity in TS and WS is intuitive as they are both measures of team quality. One would expect that the strongest teams are

also the one's that employ the most effective players. That said, both variables are statistically significant and appear to have little effect on the key independent variable, home.

The coefficient of the home variable, the key explanatory variable of the study, also remains relatively stable and robust when new variables are introduced. Each of the above specifications yield a result between five and six points-per-game. This supports the assumption that playing at home provides a significant advantage to an NBA team.

The Distance Problem

Aside from hostile crowds, cramped dressing rooms and biased referees, one would assume that travel is a primary factor in the road team disadvantage. It is then curious that the distance travelled is found to be negligible and not significant. Since this is counterintuitive, it will be investigated further and another approach is necessary in exploring the problem.

Teams were grouped into two groups - high and low travel teams - based on their cumulative travel distance over the course of the season. It was found that the median distance travelled is 70850 kms with a range of (61500, 92600). Teams were placed into high and low travel groups relative to the league median. The groups are comprised of 16 and 14 teams respectively. The groups are not of equal size because two teams - Utah and Denver - travelled an equal amount of distance above the median (71700 kms, rounded to the nearest hundred) and were both included the high travel group. The OLS regression was run in the same manner with each of the 2 groups. The results are shown in Table 3 and 4:

Table 3: Regression on Low-Travel teams

Variable	Mean	SD	Coeff	SE	t
Outcome	-0.815	13.708			
Constant			-2.494	0.721	-3.46
TS _t - TS _o	-0.959	6.934	0.678	0.075	9.037
Home	0.489	0.5	4.909	1.049	4.678
DR _t - DR _o	0.074	1.149	1.025	0.458	2.239
WS _t - WS _o	-0.968	16.872	0.052	0.011	4.698
n = 559* r ² = 0.279 F = 54.320					

* It appears the schedule makers prefer the LA Clippers over the LA Lakers, as they were required to travel 11500 less than their Staples Centre roommates.

Table 4: Regression for High-Travel teams

Variable	Mean	SD	Coeff	SE	t
Outcome	1.579	13.279			
Constant			-1.62	0.704	-2.302
TS _t - TS _o	0.658	7.458	0.609	0.068	8.92
Home	0.509	0.5	5.368	0.995	5.393
DR _t - DR _o	-0.061	1.138	0.165	0.439	0.377
WS _t - WS _o	0.691	17.527	0.032	0.012	2.78
n = 636 r ² = 0.188 F = 36.409					

In each case, when considering the t-statistics, the results yield a positive and significant home coefficient. It appears that high-travel teams experience a greater home effect than low-travel teams (5.368 points to 4.909 points). However, the relatively high standards errors for the home variable in each group suggest the coefficient may not be significant.

To test the null hypothesis that $Home_L = Home_H$, the Chow test was conducted where:

$$Sc = 175985, SL = 84941, SH = 90970, k = 5, N1 = 559, N2 = 636$$

and follows:

$$\frac{Sc - (SL + SH) / k}{(SL + SH) / (N1 + N2 - 2k)}$$

It is found that $F = 0.498$ and the null hypothesis cannot be rejected. The comparative practice is then inconclusive in determining a significant difference in home advantage for high and low travel teams.

Ashman et. al (2010) also used total distance travelled in their study of NBA wagering markets and found no significant effect. Rather, they focused much of their work on studying discrepancies in rest between home and away teams and found that travel across time zones, specifically west to east, had an adverse effect on performance. It is reasonable to conclude that weariness derived from travel, rather than the act of travel itself, is a contributing influence on performance.

Testing the Home Advantage

To test the notion that home advantage is derived from an immeasurable sense of comfort and familiarity, the concept of home team can be redefined. Instead of home being attributed to the physical location of the game, the home team will now be one that does not travel prior to the game. The

expectation is that, if the home effect is attributed to positive reinforcement from a friendly crowd and familiar conditions of a home stadium, the effect will decrease in the revised regression.

The regression was run in the same manner as the general model, and the results are shown in Table 5.

Table 5: Regression with Home being redefined as a team that did not travel prior to the game

Variable	Mean	SD	Coeff.	t
Outcome	0.464	13.535		
Constant			-0.203	-0.469
TS _t - TS _o	-0.105	7.269	0.594	11.532
Home	0.254	0.436	3.137	3.221
DR _t - DR _o	0.003	1.126	0.921	2.86
WS _t - WS _o	-0.944	34.831	0.071	6.643
D _t - D _o	23.412	1246.944	1.0x10 ⁻⁴	-0.288
n = 1195	r ² = 0.197	F=58.296	nHome = 305*	nAway = 890

*Brooklyn Nets vs. New York Knicks is considered 0 travel.

The results above show that the home effect remained positive and significant but decreased as expected. We see that under the new definition, the number of home teams decreased from 597 to 305. It is assumed that 295 teams that are now considered away teams but playing in a home arena are still experiencing the home effect. By shifting these teams to the away group, the model is effectively diluting the home effect from 5.39 points per game to 3.14. This regression does not prove the value of the home advantage effect in itself but provides reinforcement that the model is constructed accurately.

Is Home-court constant?

In each regression thus far, the home variable has been a simple dummy variable to indicate the advantage is constant. Intuitively, one would assume the home effect increases the longer a team can avoid traveling and conversely, the negative effect will compound on road trips where a team is forced to endure constant jet-lag and travel complications. It is worth considering that part of the home advantage can be attributed to travel fatigue of the road team. To test the idea, the home variable was again modified, this time to capture consecutive home/away games. In this iteration, $home^* = 1$ for a team's first home game following an away game, and $home^* = i$ for the i^{th} subsequent home game prior to the next road game. Conversely $home^* = -1$ for the first road game following a home game, and $home^* = -i$ for the i^{th} subsequent away game. All other independent variables remain the same and the OLS regression is run in the same manner as the base model. The results are shown in Table 6:

Table 6: Home redefined to reflect length of road trip or home stand.

Variable	Coeff	t
Constant	0.65	1.857
$TS_t - TS_o$	0.596	11.631
Home*	0.856	5.34
$DR_t - DR_o$	0.739	2.293
$WS_t - WS_o$	0.072	6.755
n = 1195	$r^2 = 0.205$	F = 76.764

*Home: mean = -0.064, SD = 2.263

It is found that TS (0.600 to 0.596) and WS (0.073 to 0.072) are very consistent with the findings in the base model, further reinforcing their independence from the home variable.

The home effect is again found to be positive and significant; however the value taken by the coefficient of the home variable is smaller than that in the base regression, as expected. The fact that the variable yields a sufficiently high t-ratio pays credence to the idea that the home effect can compound the longer a team can play home games without travelling. Conversely, when the team travels (i.e., a negative home value), its players will experience an increasing disadvantage for each consecutive game they play on the road.

Entine and Small (2008) attempted to address the question of the length of a road trip on the home advantage effect. By regressing differences in days of rest on game outcomes, they found that the home team experienced an additional 1.04 point advantage when isolating for cases where the road team played its second consecutive game on the road. Cases with more than two consecutive games on the road, however, did not yield significant results in their study. The findings support the assumption that road teams experience travel weariness that affects their performance on the court.

Playoffs

Home court advantage is granted as a reward during the NBA playoffs for high performance during the regular season. Playoff series are played in a best-of-seven format with the team granted home advantage given the opportunity to play a maximum of 4 games at home. The playoffs are organized with 8 teams from each conference with the best regular season record. The top 4 teams in each conference are awarded home-court in at least the first round (division winners are automatically placed in the top 4 regardless of record). In each subsequent round, the team with the best record receives the home advantage.

During the playoffs, the home-away effect becomes a very popular talking point for players, coaches, and the media. The attention on the subject is more pronounced because stakes are raised and fan enthusiasm is intensified. A seventh and final game of a series sets the stage for a dramatic winner-take-all contest. It is assumed that the home effect is maximized in games where the team's season is on the line.

To attempt to capture the effect, the base model is once again applied but including only playoff games from the 2013-2014 season. The results are as follows in Table 7:

Table 7: Regression with playoffs games exclusively.

Variable	Mean	SD	Coeff	t
Outcome	-0.652	13.067		
Constant			-2.117	-1.119
$TS_t - TS_o$	-0.119	7.174	0.516	0.922
Home	0.494	0.503	3.416	1.266
$DR_t - DR_o$	-0.067	1.031	2.577	1.935
$WS_t - WS_o$	-0.136	9.93	-0.088	-0.218
n = 89	$r^2 = 0.107$	F = 2.518	nHome = 44	nAway = 45

At first glance the results suffer from a low sample size that generates values that are not significantly different from zero. The values, with the exception of WS, have the correct signs so it is expected that increasing the sample size would produce results more consistent with regular season

findings. However, when considering the playoffs, it is seen that the problems with this model run deeper than low sample size.

The variable TS, if defined in the same way as the regular season, will not be as effective in explaining team relative strength. To recap, TS during the regular season was defined as:

$$(\text{sum of points scored in games 1 to } n-1 / n-1) - (\text{sum of points allowed in games 1 to } n-1 / n-1)$$

In the playoffs, where n is low, the high variance in TS is expected to restrict its ability to represent team strength. Half of the league's 16 playoff teams will play not more than 7 games and it is expected that their game-to-game TS is too volatile to represent their season long performance. For this, TS in the above results is simply defined as:

$$(\text{regular season avg. points scored per game}) - (\text{regular season avg. points against per game})$$

TS_{ij} is then a constant during a playoff series between team i and team j and will appear as TS_{ij} or $-(TS_{ij})$ where $-(TS_{ij}) = TS_{ji}$.

The variable DR is also problematic when applied in the same manner as the regular season. Because two teams play each other exclusively over the course of a series, their rest days are nearly always equal. The only possible discrepancy in days of rest is in the first game of the series, where the advantage typically goes to the team that won its previous series more quickly. The result of this sameness is that relative $DR = 0$ in 78 of 89 observations. Of the 11 observations where DR is not zero, the range is (4,-6). It is assumed that the low sample size, coupled with a high deviation, limits the credibility of the DR variable.

Win Shares are found not to be significant again due to a low sample dataset. Like TS, WS will appear as WS_{ij} or $-(WS_{ij})$ depending on the observed team in a particular game. In the playoffs, there are

far less different match ups than in the regular season, and because two teams will play one another repeatedly, there is less variation of WS in the data.

The first and most obvious solution for remedying the issue would be to compile playoff data from at least the past 14 seasons in order to replicate the sample size of one regular season¹. Since the model utilizes TS and WS in relative terms, it is not likely that rule changes or league wide performance trends over time will have a significant influence on the results.

The sample size limitation of this model appears fairly straightforward in this case, but what is worth discussing are some determining factors that are magnified during the playoffs. Since two teams play each other anywhere from 4 to 7 consecutive times, coaches, scouts and players have more time to study the opposing team and develop more informed strategic decisions compared to the regular season when opponent turnover is rapid. The effectiveness in an organization's ability to adapt strategic decisions unique to a particular opponent will play a pronounced role during the playoffs. In any study focusing on playoff outcomes, one would require a method to control "coaching skill" to capture this effect. It is expected that this practice would be very difficult to achieve since coaching is not directly measured and fans and researchers are not privy to many aspects of team strategy.

A fan of any sport has likely heard the adage that "defense wins championships." It is often preached as rhetoric come playoff time by both players and the media. The expectation in NBA basketball is the game in the playoffs slows down and younger, faster teams are not able to simply outscore their way to the championship. It is often assumed that veteran leadership and experience provide invaluable contributions during the playoffs. To test these long-standing assumptions, it is likely any future study of the playoffs is required to control for team experience, pace, and defensive efficiency.

¹ There have been averages of 84 playoffs games per year over the past 5 seasons, per basketball-reference.com game logs.

Price, Remer, and Stone (2012) tested the hypothesis that referees have a bias toward teams that are behind in the playoff series under the premise that the league has an incentive to lengthen series' to maximize profits. They found that for each game a team is down (up) in a series, it gains (losses) a 3.4% advantage in turnovers. It is difficult to say that this effect reflects league executives' greed or is enough to satisfy the losing team's fans cries of conspiracy, but there is significant evidence of referee bias exclusive to the playoffs.

Further investigation of the home effect during the playoffs likely deserves its own model that is better designed to address these potential shortcomings. The methods presented in this paper are suited to explore the NBA regular season; however the inherent differences in playoff scheduling and performance make it sub-optimal to simply apply the same methods to the playoffs.

5. Conclusion

The primary motivation for this study was to capture a point-estimate for the NBA home-court advantage and the finding, between 5 and 6 points per game, is greater than that of Entine and Small (2008) estimate of 3.24 points. Where this paper is intended to differ is in the inclusion of the advanced player metric Win Shares. In each of the regressions presented, it was found that Win Shares were significantly different than zero, indicating that the statistic is a reliable evaluation method and it does belong in analysis such as this. That is not to say Win Shares is perfect by any means. Those who study basketball metrics continue to refine evaluation methods in pursuit of an all-encompassing player rater.

Most advanced metrics - like win shares - are primarily built of offensive production since typically when it comes to scoring points, the ends justify the means. That is, by studying player field

goals, free throws, and assists, one can reasonably deduce a player's offensive value (Bornn et al., 2015). NBA analysts, however, have long struggled to understand player defensive contributions beyond simple, results-based statistics such as steals, blocks, rebounds and fouls. Professional defenses are highly coordinated, synchronized team efforts built around a system of concepts like switching, hedging, and double teaming. An efficient defense is one where five players move and communicate as one group. When breakdowns occur, it is obvious to the viewer when witnessing an easy basket. However, the player unfortunate enough to be on the receiving end of a dunk is not always the one responsible for surrendering the basket.

Bornn et al. (2015) introduced the concept of "counterpoints" as a defensive metric of individual points against at the 2015 MIT Sloan Sports Analytics Conference. They believe they have leveraged SportVU player tracking to establish a counterbalance to the existing player metrics heavily influenced by offensive production. With these recent advancements on defensive statistics it appears that win shares, although shown to hold a place in this type of study, holds room for improvement. Counterpoints are a new innovation that is not yet widely accepted so data remains unavailable to the public. As NBA statisticians continue to refine player evaluation ratings, future attempts to revisit the home advantage effect should be modified to reflect the best player rater available at the time.

It is probably unlikely that NBA coaches and players would allow to a point estimate of home advantage to change their strategic approach. The NBA awards the winner as the team who scores the most points, regardless of the margin of victory. It is unlikely that coaches would invest in the notion that their team is "spotted" 5 points by virtue of playing at home. It is assumed that most would take the stance that every point must be earned, and they wouldn't be wrong in believing so. For those under pressure to perform in the arena, this estimate may not be so useful. It does, however, hold potential for anyone interested in studying league-wide trends.

The estimate is likely to be of great use to game handicappers in sports betting. Although most fans acknowledge the home advantage and are wise to consider it in handicapping, the magnitude of the effect is not necessarily common knowledge or readily shared by sports books. Gandar, Zuber, and Lamb (2001) found no evidence of mispricing the home team advantage in NBA betting markets. They concluded that the market is, generally, efficient.

This paper does not claim that its findings are capable of exploiting market inefficiencies. These findings, however, may be useful aids in a gambling prospector's decision making process. For analysts, the estimates presented may be applied to future studies on the basketball betting markets.

Pollard and Pollard (2005) presented evidence of a declining home advantage by inspecting the home team's winning percentage season-by-season since the league's inauguration. Expressing the effect in terms of win percentage fails to explain its true value. Rather than simply observing home team's win/loss outcomes over time, studying changes in home advantage over time can be improved upon by replicating the practices in this paper on a season-to-season basis.

As NBA analytics move forward with increasingly sophisticated team and player evaluation methods, the studies on the influential factors on outcomes should be equally progressive. Rather than simply observing home and away win/loss records as some studies have in the past, this paper makes strides in developing understanding of the magnitude of discrepancy in home and away performance. The findings of the study are robust and the framework is designed to seamlessly evolve with the landscape of NBA analytics.

Appendix 1 - Calculating Win Shares

Win Shares are a player evaluation metric first developed by Bill James (2002) in his book *Win Shares*. Justin Kubatko, founder of www.basketball-reference.com and former statistical consultant for the Portland Trail Blazers, then applied James' concepts to player evaluation methods outlined in the Dean Oliver 2004 book *Basketball on Paper* and further in an Oliver et. al (2007) paper. The purpose of the statistic is to capture an individual's player's contribution to his team's win total. It then reasons that the sum of player win shares on a team will be equal that team's win total. It was found that since the 1962-63 season, the average error between team wins and sum of player win shares is 2.74 wins.

To illustrate Win Shares and their application to this paper, consider the 2013-2014 Oklahoma City Thunder and league MVP Kevin Durant. The Thunder finished the season with a record of 59-23 - the second best win total in the league - and their players' cumulative Win Shares was 62.5. Kevin Durant led the league in this statistic with 19.2. For any game that Durant did not play, the Thunder's cumulative WS is 43.3, which effectively make Oklahoma City a league average team.

Calculating Win Shares is described below, as per basketball-reference.com:

First one must calculate the player's points produced and offensive possessions.

Calculate marginal offense for each player. Marginal offense is equal to $(\text{points produced}) - 0.92 * (\text{league points per possession}) * (\text{offensive possessions})$. Note that this formula may produce a negative result for some players.

Calculate marginal points per win. Marginal points per win reduces to $0.32 * (\text{league points per game}) * ((\text{team pace}) / (\text{league pace}))$.

Credit Offensive Win Shares to the players. Offensive Win Shares are credited using the following formula: (marginal offense) / (marginal points per win).

Calculate the Defensive Rating for each player².

Calculate marginal defense for each player. Marginal defense is equal to (player minutes played / team minutes played) * (team defensive possessions) * (1.08 * (league points per possession) - ((Defensive Rating) / 100)). Note that this formula may produce a negative result for some players.

Calculate marginal points per win. Marginal points per win reduces to $0.32 * (\text{league points per game}) * ((\text{team pace}) / (\text{league pace}))$.

Credit Defensive Win Shares to the players. Defensive Win Shares are credited using the following formula: (marginal defense) / (marginal points per win).

Add Offensive and Defensive Win Shares to obtain the player's Win Shares.

² These calculations are quite labour intensive and outlined in detail in the appendix of Oliver's *Basketball on Paper*, 2004 as well as *A Starting Point for Analyzing Basketball Statistics*, *Journal of Quantitative Analysis in Sport* (2007).

References

- Ashman, T., Bowman, R. & Lambrinos, J. (2010) "The Role of Fatigue in NBA Wagering Markets: The Surprising 'Home Disadvantage Situation.'" *Journal of Sports Economics* 11(6), 602-613.
- Bornn, L., Franks, A., Goldsberry, K. & Miller, A. (2015) "Counterpoints: Advanced Defensive Metrics for NBA Basketball," MIT Sloan Sports Analytics Conference.
- Bray, S., Carron, A. & Loughhead, T. (2005) "The Home Advantage in Sport Competitions: Courneya and Carron's (1992) Conceptual Framework a Decade Later." *Journal of Sports Sciences* 23(4), 395-407.
- Courneya, K., & Carron, A. (1992) "Sport Psychology, The Home Advantage in Sport Competitions: A Literature Review." *Journal of Sport and Exercise Psychology* 14(1), 13-27.
- Doyle, J., & Leard, B. (2011) "The Effect of Home Advantage, Momentum, and Fighting on Winning in the National Hockey League." *Journal of Sports Economics* 12(5), 538-560.
- Entine, O. & Small, D. (2008) "The Role of Rest in the NBA Home-Court Advantage," *Journal of Quantitative Analysis in Sports*, 4(2), 1106-1115.
- Gandar, J. M., Zuber, R. A. & Lamb, R. (2001) "The home field advantage revisited: A search for the bias in other sports betting markets." *Journal of Economics and Business* 53, 439-453.
- Greer, D.L. (1983) "Spectator booing and the home advantage: a study of social influence in the basketball arena." *Social Psychology Quarterly* 46, 252-261.
- James, B., & Henzler J. (2002) *Win Shares*, Stats Inc., Chicago, Illinois.
- James, B. (1982) *The Bill James Baseball Abstract*, Ballantine Books, New York, New York.
- Jurkovic T. (1983) "Collegiate basketball players' perceptions of the home advantage [thesis]." Bowling Green (OH): Bowling Green State University.
- Lewis, M. (2003) *Moneyball: The Art of Winning and Unfair Game*, W.W. Norton and Company, New York, New York.
- Nevill, A., & Holder, R., (1999) "Home Advantage in Sport: An Overview of Studies on the Advantage of Playing at Home." *Sports Medicine* 28(4), 221-236.
- Oliver, D. (2004) *Basketball on Paper: Rules and Tools for Performance Analysis*, Potomac Books Inc., Dulles, Virginia.
- Pollard, G. & Pollard, R. (2005) "Long-Term Trends in Home Advantage in Professional Team Sports in North America and England (1876-2003)." *Journal of Sports Science* 23(4), 337-350.
- Pollard, R. (1986) "Home advantage in soccer: a retrospective analysis." *Journal of Sport Science* 4, 237-248.

Pollard, R. (2002) "Evidence of a reduced home advantage when a team moves to a new stadium." *Journal of Sports Sciences* 20, 969-973.

Price, J., Remer, M. & Stone, D. (2012) "Subperfect Game: Profitable Biases of NBA Referees." *Journal of Economics & Management Strategy* 21(1), 271-300.

Puerzer, R. (2002) "From Scientific Baseball to Sabermetrics: Professional Baseball as a Reflection of Engineering and Management in Society." *Journal of Baseball History and Culture* 11(1), 34-48.

Shmanske, S. & Lowenthal, F. (2009) "Home Field (Dis)Advantage and the "Last-Ups" Effect." *Journal of Baseball History and Culture* 18(1), 130-146.

Swartz, T. & Arce, A. (2014) "New Insights Involving Home Team Advantage." *International Journal of Sports Science and Coaching* 9(4), 681.

Wang, W., Johnston, R. & Jones, K. (2011) "Home Advantage in American College Football Games: A Multilevel Approach." *Journal of Quantitative Analysis in Sports* 7(3), 1328.