No referee bias in the NBA: New evidence with leagues' assessment data

Christian Deutscher*
Bielefeld University, Bielefeld, Germany

Abstract. Beginning March 2015, the National Basketball Association (NBA) started the public assessment of officiated events in close game situations, where teams are within five points with two or less minutes to play. This ex-post league evaluation of officials' actions allows for a much improved analysis of referee biases like the home bias, preferential treatment of superstars, race and favoritism towards losing teams. Instead of relying on statistical frequency of calls and devoting it to biased decision making, in-game decision making is matched with reviewed broadcasting video in this paper. The empirical analysis for 113 games and 1229 total calls finds no support of referee bias in foul calling.

Keywords: Referee bias, favoritism, social pressure, discrimination, superstars, NBA

1. Introduction

Referees in the National Basketball Association (NBA) are hired by the league to judge games impartially. They evaluate in game situations subjectively and are potentially prone to biases that are not in line with the league's interest.[1] These biases of judgment by the referee can stem from personal preferences towards certain players or teams. Social payoffs in form of home fans applauding for calls in their teams' favor can serve as another kind of non-monetary reward. In most recent years, cases of bribing made the press, like the 2007 NBA betting scandal surrounding former referee Tim Donaghy or the 2005 Bundesliga soccer scandal centered on former referee Robert Hoyzer. Both the Donaghy case and the Hoyzer case resulted in criminal proceedings, evidencing the overlap between referee bias and potential legal issues that may result.

To lessen favoritism, leagues have various instruments to implement unbiased decision making, where the most powerful and presumably most costly tool is a monitoring system to supervise referees. Here, leagues evaluate the performance by their referees to tie chances for reappointment and promotion to the proper and impartial calling of games. Hence, financial incentives by the NBA aim at unbiased game calling by the referees. What remains is a potential trade-off for the referees, who face contradicting expectations from their employer, fans, and third parties, each with individual interest in the game outcome. Given the literature on referee bias in sports, impartial decision making by referees should not be taken for granted. The literature on referee biases report favoritism towards home teams, players of the referees' ethnicity, losing teams, and others.[2]

This paper presents a novel approach in analyzing discrepancies between actual calls and the judgment of these foul situations by the league. For the first time, differences in the assessment of game situations by the employer (NBA) and employees (referees) are

*Corresponding author: Christian Deutscher, Bielefeld University, Postfach 100 131, 33501 Bielefeld, Germany. Tel.: +49 521 106 2006; E-mail: christian.deutscher@uni-bielefeld.de.

[1]Note that this is a classic example of a principal agent problem in economics, where the agent (in this case the referee) has some room for discretion in his decision making that can be against the interest of the principal (the NBA).

[2]See Dohmen and Sauermann (2015) for an overview of the literature.

compared on a call-by-call basis. This has the advantage of a very precise measurement of calls that can be broken down to the individual player committing a foul. Given that player characteristics like origin or star status are common knowledge, it opens the possibility to devote biased decision making by referees to these player specifics. Previous investigations relied on statistical frequency of calls and devoted peculiarities to biased decision making, potentially mixing biased decision making with actual differences in behavior by the players or teams.

The next section presents a review of the literature on referee biases. Section 3 describes the league assessment data for foul calls as well as estimation results for potential referee biases. Section 4 summarizes the results and provides a short discussion.

2. Referee bias in the literature

Referees have the task to evaluate in game situations impartially and make decisions within a very short period of time. Yet, there is room for subjective interpretation of situations and a chance for biased decision making. Previous work on referee bias concentrates on statistical differences in frequencies of decisions. Home bias is presumably the most often analyzed type of bias in referee decision making. Referees let their desire for social approval impact their decision making and account for at least some of the home court advantage (Moskowitz & Wertheim, 2011). In soccer, the stoppage time is a discrete decision by the referee and multiple studies provide support for additional stoppage time at the end of games if the home team is trailing (Sutter & Kocher, 2004; Garicano et al., 2005; Scoppa, 2008). Similar findings favoring home teams are confirmed for penalties as well as yellow and red cards (Boyko et al., 2007; Dohmen, 2008; Buraimo, Forrest, & Simmons, 2010; Dawson & Dobson, 2010). For college basketball, Anderson and Pierce (2009) find more fouls to be called against away teams, while Price, Remer and Stone (2012) find discretionary decision making of turnover judgment to be in favor of home teams in the NBA. Superstar players are allegedly treated different than other players. Evidence of this is rare in the NBA. Caudill et al. (2014) find NBA All-Stars to be rewarded preferential treatment by the referees, with an additional 0.32 free throw attempts per minute in crucial game situations. Price and Wolfers (2010) probably provided the most contro-

versial study regarding referee bias in the NBA, as they find more personal fouls to be called against players by opposite-race referee crews. Subsequent to extensive media coverage after the release of the results, Pope, Price and Wolfers (2013) find this discrimination to vanish, possibly as a result of increased referee awareness. Little is known about referee bias towards the underdog in basketball. For soccer, Buraimo, Forrest and Simmons (2010) find that underdogs are less likely to receive yellow cards if they play at home. On the other hand, Dawson et al. (2007) determine the number disciplinary sanctions to be larger for underdogs compared to favorites.

3. Data, descriptive statistics and results

Studies mentioned in the previous section share the predicate of analyzing statistical frequency of calls, rather than analyzing call-by-call. Conclusions in the literature have been drawn without knowledge if referee decisions were correct, but simply on how often they occur. The data at hand for this paper adds significant value as it has detailed information on the correctness of foul calls as well as material non-calls that are called (or not called). Information is available for crucial games situations at the end of NBA games. Every call is reviewed by a senior referee manager or basketball operations manager and published online the day after the game is played. Information is available on the official website of the NBA at www.official.nba.com.

In this paper, calls are analyzed for every close game played during the 2014-15 regular season after March 1st 2015. The NBA refers games to being close when no team is ahead by more than five points with two minutes or less to play, or overtime (Deutscher, Frick, & Prinz, 2013). Out of the 356 regular season games during the period under observation, 113 fit this criterion. The data include 1229 calls and material non-calls that can be classified as displayed in Table 1[3].

Find the frequencies for actual decisions by the referee (foul called and no foul called) as well as the assessment by the league (foul committed and no foul committed) in Table 1. 496 out of 619 fouls identifies by the league are correctly called by the referees

---

[3]Note that only personal fouls are included in the sample while technical fouls as well as other violations and reviewed decisions are excluded as they fundamentally differ from foul calls and are observed very unfrequently.

(80.1 percent) while 593 of 610 no-foul situations identified by the league are correctly not called by the referees (97.1 percent). Due to the diminishing percentage of incorrect calls in no foul situations, these cases are dismissed from the analysis as the variation between players is too low. The following analysis focuses on fouls as assessed by NBA referees that are either called or not called by the referee. Here, the dependent dummy variable *correct call* indicates if the referee called the foul.

To analyze potential biases described in the literature review, additional data for every incident documented by the league has been added to the data from the website www.basketball-reference.com. The aim here is to systematically control for biases known from the literature, with the virtue of not having to rely on statistical frequencies of calls, but analyzing call-by-call decision making. Additionally, this is the first paper that combines the usage of different biases into one comprehensive approach. Control variables for often identified biases include information if the call is against the home (*home*) or away team (e.g. Anderson & Pierce, 2009), and if the committing and fouled players can be referred to as a superstar (*Star*) or not (*NonStar*) (e.g. Caudill et al., 2014). To control for a potential own-nationality bias (Pope & Pope, 2015), this paper distinguishes if the players' origin is within the United States (*US*) or not (*NonUS*) as well as the underdog/favorite (*favorite*) status of their teams (e.g. Dawson et al., 2007). I define players as superstars if their number of appearances in NBA all-star games is at least one standard deviation above the average value for the full sample (Frick, 2001). For this paper, a player needs to appear in at least three NBA all-star games to be referred to as a superstar, a requirement 6.6 percent of the players in the sample meet. Teams are classified as underdogs if their probability to win was determined to be below 50 percent by the bookmaker prior to the game. Betting odds were drawn from the website betexplorer.com.

Further control variables include the number of seconds left in the game as well as crowd presence in the arena. While seconds left (*secondsleft*) to play in the game serve as an indicator for importance of a situation and the pressure on the decision by the referee, crowd presence (*crowd presence*) is measured as the percentage of tickets sold to display the crowd presence in the arena and possible social payoff to decisions in favor of the home team (Dohmen, 2008). As NBA arenas are very similar in their architecture, the necessity of further control variables is limited (Deutscher, 2011).

While every bias towards the home team would mean a negative bias against the away team, the introduction of superstar status and origin of players is more complex as, for example, the player committing a foul as well as the player being fouled could be a superstar. For superstars status as well as the origin of the players, four possible constellations are possible for fouls, as the player committing the foul as well as the player being fouled can fit or not fit the criteria superstar or US origin. All possible constellations are displayed in Table 2a and 2b and label the respective dummy variables and number of observations to be included in the empirical analysis. This labeling serves as a novel approach, since this paper is the first to distinguish between players committing fouls and players being fouled (foul committed and no foul committed).

Given these classification of fouls, Table 3 offers descriptive statistics for 619 fouls that were either correct calls or incorrect non calls. 19.7 percent of the fouls involved at least one superstar, while 38.1 percent of the fouls involved at least one foreign player.

To test foul calls for referee biases, we turn the attention to our dependent variable *correct call*. Its nature as a dummy variable suggests to apply a logit approach (Cox, 1958). The independent control

Table 1

League assessment information and referee decisions on foul calls

	foul called	no foul called
foul committed	correct call ($N = 496$)	incorrect non call ($N = 123$)
no foul committed	incorrect call ($N = 17$)	correct non call ($N = 593$)

Table 2a

Labeling of fouls in accordance to superstar status ($N = 619$)

		Player committing foul	
		Star	Non-Star
Player	Star	Star vs. Star ($N = 14$)	Non-Star vs. Star ($N = 70$)
fouled	Non-Star	Star vs. Non-Star ($N = 38$)	Non-Star-vs. Non-Star ($N = 497$)

Table 2b
Labeling of fouls and players' origin ($N = 619$)

		Player committing foul	
		US	Non-US
Player	US	US vs. US ($N = 383$)	Non-US vs. US ($N = 108$)
fouled	Non-US	US vs. Non-US ($N = 89$)	Non-US vs. Non-US ($N = 39$)

Table 3
Descriptive statistics for fouls

Variable	Mean	Min.	Max.
dependent variable			
Correct call	0.80	0	1
independent variables			
Home	0.48	0	1
Star vs. Star	0.02	0	1
Non-Star vs. Star	0.11	0	1
Star vs. Non-Star	0.06	0	1
Non-Star vs. Non-Star	0.80	0	1
US vs. US	0.62	0	1
Non-US vs. US	0.17	0	1
US vs. Non-US	0.14	0	1
Non-US vs. Non-US	0.06	0	1
Favorite	0.45	0	1
Secondsleft	64.18	0.2	291
Crowd Presence	0.93	0.57	1.03

variables account for potential referee biases towards home teams, superstar players, players with US origin and favorite teams while seconds left to play and the attendance serve as further control variables. Table 4 displays the logit estimations for correct calls in crucial game situations, where no team is ahead by more than five points and there are 2 minutes or less to play or overtime. While Model 1 estimates the impact of the most common referee bias (home bias), subsequent models include further referee biases described in the previous sections.

Results are very robust throughout all models. This data provides no support for home bias, contradictory to the vast majority of the literature. For an average value of 19.9 percent missed calls, no subgroup except for underdogs exhibits a value that is significantly different with 90% confidence. Concerning favoritism towards superstar (which would be expected in "Star vs Non- Star" or "Non-Star vs. Star" situations) no bias can be found compared to "neutral" foul situations (where a non-star fouls a non-star). Fouls where either a player from the US fouls or is fouled by a non-US player also provide no evidence of biased referee decision making. Compared to fouls without any player from the US, no systematic bias is detected by the estimations. Last NBA referees show a weak preference towards underdog teams. Control variables capturing

the attendance and time left to play are not significant in any model.

Throughout all models, referee bias of the type tested in this paper appears to be largely non- existent in the NBA in crucial game situations. Reasons can be manifold: For referees, financial incentives to be achieved by reappointments in the future can serve as an explanation for the results. If the league punishes biased decision making, referees have an incentive for impartial behavior. Second, the NBA could fear bad press in case biased referee decision making becomes publically known. Referee bias as documented in academia by Price and Wolfers (2010) is not supported for later seasons (Pope, Price, & Wolfers, 2013). While no official statement by the NBA documents changes related to the results published by Price and Wolfers (2010), Price and Wolfers (2013) is at least suggestive of an improvement in referee training or monitoring. Furthermore, the sample size is a potential problem for the estimation.

Using information on assessment of calls by the league itself comprises the potential problem of bias judgment. If the person judging the call ex-post is biased the results in the estimations above would provide no support for biased judgment by the referees. In economic terms, the question "Who monitors the monitor?" remains. The data provides further limitations to be mentioned. First, there is only information on calls in crucial situations of close games. Referee bias could prevail in foul calling earlier in games or in games decided early. Second, monitoring by the league would fail if referee decision making is only evaluated by the league for predictable game situations. Third, the NBA allows for video revisions late during games to reduce the probability of bad referee decisions. This again reduces the probability of bad calls as certain calls can be revised.

4. Conclusion

Professional and unbiased decision making by referees is crucial to the integrity of any professional sports league and the avoidance of problematic legal

Table 4
Logit estimation for correct foul calls in close game situations

	Model 1 Coefficient	Model 2 Coefficient	Model 3 Coefficient	Model 4 Coefficient
Home	0.21	0.21	0.21	0.07
	(1.05)+	(1.01)+	(1.03)+	(0.34)+
Star vs. Star		−0.68	−0.97	−0.96
		(−1.18)+	(−1.66)*	(−1.63)+
Non-Star vs. Star		0.52	0.51	0.43
		(1.39)+	(1.34)+	(1.11)+
Star vs. Non-Star		−0.00	−0.14	−0.09
		(−0.00)+	(0.00)+	(−0.21)+
Non-Star vs. Non-Star		Reference	Reference	Reference
US vs. US			0.46	0.43
			(1.12)+	(1.05)+
Non-US vs. US			−0.40	−0.38
			(−0.91)+	(−0.85)
US vs. Non-US			−0.23	−0.25
			(−0.50)+	(−0.54)
Non-US vs. Non-US			Reference	Reference
Favorite				−0.43
				(−1.95)*
Secondsleft	−0.00	−0.00	−0.00	−0.00
	(−1.47)+	(−1.22)+	(−1.47)+	(−1.07)+
Attendance	−0.70	−0.69	−0.70	−0.49
	(−0.71)+	(−0.69)+	(−0.71)+	(−0.47)+
N	619	619	619	619
Pseudo R²	0.01	0.01	0.03	0.04

Note: z-values in brackets; *$p < 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$; +n.s.

issues. As money spent on tickets, commercials, and broadcasting rights increases rapidly, the NBA requires professional training and impartial employees. For late-game foul calls assessed by the league's referees, this paper does not find support of certain referee biases documented in the literature. Future research should take advantage of the richness of the data to analyze playoffs separately from regular season games as the pressure is at its highest for players and referees during the post season. To analyze biases on an individual referee basis, more observations and games are needed. Given additional data, the possibility arises to account for differences in liabilities between referee crews. This would help the league to implement individual training for referees, contingent on their vulnerability for biased decision making.

References

Anderson, K.J., Pierce, D.A., 2009. The Effect of Foul Differential on Subsequent Foul Calls in NCAA Basketball. Journal of Sports Sciences. 27(7), 687–694.

Boyko, R.H., Boyko, A.R., Boyko, M.G., 2007. Referee Bias Contributes to Home Advantage in English Premiership Football. Journal of Sports Sciences. 25(11), 1185–1194.

Buraimo, B., Forrest, D., Simmons, R., 2010. The 12th Man? Refereeing Bias in English and German soccer. Journal of the Royal Statistical Society: Series A (Statistics in Society). 173(2), 431–449.

Caudill, S.B., Mixon J.R., F.G., Wallace, S., 2014. Life on the Red Carpet: Star Players and Referee Bias in the National Basketball Association. International Journal of the Economics of Business. 21(2), 245–253.

Cox, D.R., 1958. The regression analysis of binary sequences. Journal of the Royal Statistical Society. Series. 215–242.

Dawson, P., Dobson, S., 2010. The Influence of Social Pressure and Nationality on Individual Decisions: Evidence from the Behaviour of Referees. Journal of Economic Psychology. 31(2), 181–191.

Dawson, P., Dobson, S., Goddard, J., Wilson, J., 2007 Are football referees really biased and inconsistent? Evidence on the incidence of disciplinary sanction in the English Premier League. Journal of the Royal Statistical Society: Series A (Statistics in Society) 170(1), 231–250.

Deutscher, C., 2011: Productivity and New Audiences: Empirical Evidence from Professional Basketball. Journal of Sports Economics 12(3), 391–403.

Deutscher, C., Frick B., Prinz, J., 2013. Performance under pressure: Estimating the returns to mental strength in professional basketball. European Sport Management Quarterly. 13(2), 216–231.

Dohmen, T.J., 2008. The Influence of Social Forces: Evidence from the Behavior of Football Referees. Economic Inquiry. 46(3), 411–424.

Dohmen, T., Sauermann, J., 2015. Referee Bias. Journal of Economic Surveys, forthcoming.

Frick, B. (2001). Die Einkommen von,, Superstars" und,, Wasserträgern "im professionellen Team-Sport – Ökonomische Analyse und empirische Befunde. Zeitschrift für Betriebswirtschaft. 71(6), 701–720.

Garicano, L., Palacios-Huerta, I., Prendergast, C. 2005. Favoritism under Social Pressure. Review of Economics and Statistics. 87(2), 208–216.

Moskowitz, T., Wertheim, L.J., 2011. Scorecasting: The Hidden Influences Behind How Sports are Played and Games are Won. Crown Archetype.

Pope, B.R., Pope, N.G., 2015. Own-Nationality Bias: Evidence from UEFA Champions League Football Referees. Economic Inquiry. 53(2), 1292–1304.

Pope, D.G., Price, J., Wolfers, J., 2013. Awareness reduces racial bias (No. w19765). *National Bureau of Economic Research*.

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

Price, J., Wolfers, J., 2010. Racial Discrimination among NBA Referees. Quarterly Journal of Economics. 125(4), 1859–1887.

Scoppa, V., 2008. Are Subjective Evaluations Biased by Social Factors or Connections? An Econometric Analysis of Soccer Referee Decisions. Empirical Economics. 35(1), 123–140.

Sutter, M., & Kocher, M.G., (2004). Favoritism of Agents–the Case of Referees' Home Bias. Journal of Economic Psychology. 25(4), 461–469.