

This article was downloaded by: [Democritus University of Thrace]

On: 16 December 2012, At: 03:23

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



European Sport Management Quarterly

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/resm20>

Payback Calls: A Starting Point for Measuring Basketball Referee Bias and Impact on Team Performance

Ryan M. Rodenberg^a & Choong Hoon Lim^a

^a Indiana University, Bloomington, Indiana, USA

Version of record first published: 20 Nov 2009.

To cite this article: Ryan M. Rodenberg & Choong Hoon Lim (2009): Payback Calls: A Starting Point for Measuring Basketball Referee Bias and Impact on Team Performance, *European Sport Management Quarterly*, 9:4, 375-387

To link to this article: <http://dx.doi.org/10.1080/16184740903331853>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

ARTICLE

Payback Calls: A Starting Point for Measuring Basketball Referee Bias and Impact on Team Performance

RYAN M. RODENBERG & CHOONG HOON LIM

Indiana University, Bloomington, Indiana, USA

ABSTRACT Recent scandals in sports have (re-)emphasized the need for targeted monitoring that is legal, reasonable and effective. The National Basketball Association (NBA) provides an ideal environment to measure the effect of individual referees on team performance and non-conclusively test for possible bias by referees against league teams and affiliated individuals. In the course of analyzing 654 games and 77 referees over seven NBA seasons, we find that no NBA referee had a significant adverse effect on team performance or exhibited bias against the Dallas Mavericks when considering all games (regular season and playoffs). However, when analyzing only the 80 playoff games involving the team, we find one example of an NBA referee having a significantly adverse effect on team performance. Retribution theory is used to explain the possibility of such a *prima facie* finding. Nevertheless, given our use of non-conclusive indirect detection methods, such a finding merely gives rise to a rebuttable presumption.

Introduction

We know that the NBA's success hinges on the integrity of our sport and on competition that allows teams to win based on their own skill and performance. We expect nothing less from our referees than the highest level of accuracy, professionalism, and integrity (David Stern).¹

Price and Wolfers (2007) showed how data-rich professional basketball in the National Basketball Association (NBA) provides an ideal environment to evaluate whether decision makers charged with the duty to be objective can maintain impartiality. This paper aims to measure the impact of individual

Correspondence Address: Ryan M. Rodenberg, Indiana University, 1025 E. 7th St, HPER #112, Bloomington, Indiana 47405, USA. Fax: 812-855-3193; Tel.: 812-855-4738; Email: rmrodenb@indiana.edu

referees on team performance and non-conclusively test whether one or more NBA referees are biased against the Dallas Mavericks, an NBA team with a majority owner who has a history of making public comments questioning the competency and impartiality of NBA referees. Pedowitz (2008, p. 57) highlighted the importance of such research when he declared that “[basketball] referee bias is a threat to the integrity of the game”. In addition to its primary purpose, this paper is secondarily aimed at offering an analytical model that could be used by various sports governing bodies in their monitoring of game officials. Such an effort is important as professional (and amateur) sports leagues seek to preserve the integrity of their sporting contests after dealing with recent highly publicized scandals, corruption and criminality that have included, but are not limited to, gambling, doping, fraud, point shaving, illegal immigration, employee tampering, espionage, interference with contractual relations, and dog fighting.

Team performance has been analyzed across a myriad of different sports. Espitia-Escuer and Garcia-Cebrian (2006, p. 1021) presented a sample of papers “dedicated to evaluating team and sports performance using economic tools”. For example, economic aspects of team performance have been investigated in Major League Baseball (Scully, 1974; Kahn, 1993), the National Football League (Hadley, Poitras, Ruggiero, & Knowles, 2000) and English Premier League soccer (Carmichael, Thomas, & Ward, 2000). Specific to the NBA, team performance has been analyzed as a function of coaching (Berri, Leeds, Leeds, & Mondello, 2009) and wages (Berri & Jewell, 2004). This paper adds to the literature by exploring basketball team performance *vis-à-vis* another variable, on-court referees.

This paper begins by introducing our model and discussing its theoretical underpinnings. We focus on the Dallas Mavericks and its majority owner Mark Cuban, who recently opined that: “There’s not ten players on the court, there’s 13 ... and three of them determine about 80% of what happens out there” (Windhorst, 2009). We also emphasize relevant past conduct by NBA referees. We then describe our dataset and analyze the adverse effect specific referees may have had on the team’s performance, with a particular enquiry into whether any referee exhibits bias during more meaningful playoff games. We adopt retribution theory to explain our possible bias-related findings. We conclude by explaining why this analysis is merely the “starting point” for making any inferences and discussing the limitations and potential extensions of our work.

Model and Theory

This paper measures the effect of NBA referees on the Dallas Mavericks’ team performance by calculating three different z-scores for every referee who officiated a Dallas Mavericks game during a seven-year time period. Harris (1998) posited that z-scores are the most widely used and generally accepted kind of standardized scores. Calculating z-scores for each referee allows for a comparison of observations from different distributions. An

initial z-score for each referee was determined by using the team ratings created by Jeff Sagarin and published in the *USA Today*. Among other things, such ratings are used to estimate expected margins of victory in a variety of sports, including professional basketball. Bernhardt and Heston (2009) used the Sagarin rating system in the context of making predictions about the outcome of sports contests for the purpose of detecting point shaving in college basketball. The algorithms and methods used by Sagarin are similar to those described by Harville (1977; 1980) and Dare and MacDonald (1996). Winston (2007) explained how ratings for basketball, baseball, football and hockey teams similar to the proprietary ones developed by Sagarin can be generated using Microsoft Excel's Solver feature. Bernhardt and Heston (2009, p. 9) concluded that Sagarin "team ratings are a very good proxy for expected winning margins". The formula for individual basketball referee z-scores based on Sagarin's predicted margin of victory ($zSAG$) is as follows:

$$zSAG = \mu_{\text{residual per ref}} \div \left(\sigma_{\text{residuals all games}} \div \sqrt{\text{games per ref}} \right) \quad (1)$$

where $\mu_{\text{residual per ref}}$ is the average "Sagarin prediction" residual for each game that a specific referee worked, $\sigma_{\text{residuals all games}}$ is the standard deviation of the "Sagarin prediction" residuals of all games in the dataset and $\sqrt{\text{games per ref}}$ is the square root of the number of games officiated by each referee.

A second z-score using the Sagarin ratings was determined by analyzing the Mavericks' actual wins and losses from 2001–2002 to 2007–2008 *vis-à-vis* Sagarin's prediction of the game outcome, without regard to the (actual or predicted) margin of victory or defeat. The formula for individual basketball referee z-scores based on Sagarin's predicted win–loss outcome (zWL) is as follows:

$$zWL = \left(X_{\text{actual wins per ref}} - \mu_{\text{Sagarin predicted wins per ref}} \right) \div \sigma_{\text{actual wins per ref}} \quad (2)$$

where $X_{\text{actual wins per ref}}$ is the actual number of games won by the Dallas Mavericks when a specific referee was working, $\mu_{\text{Sagarin predicted wins per ref}}$ is the percentage of time the Mavericks are expected to win the game officiated by a specific referee assuming normal distribution, and $\sigma_{\text{actual wins per ref}}$ is the standard deviation of games in which the Mavericks prevailed when a certain referee was officiating.

A third z-score for each referee was calculated by using point spreads published by Covers.com. The actual margin of victory was compared to the point spread in every game. Larson, Price, and Wolfers (2008) used point spreads from Covers.com in the course of analyzing the betting market implications stemming from Price and Wolfers's (2007) finding of racial bias by NBA referees. Likewise, Schnytzer and Weinberg (2004) employed Covers.com's point spreads in their work determining whether the NBA betting market is efficient. Basketball point spreads are used primarily in the gambling industry to balance the level of wagering between the team favored to win and the "underdog". Dare and MacDonald (1996, p. 295) explained

that bookmakers try to set the point spreads so that “equivalent total dollar amounts are wagered on each team”. Oorlog (1995) provided a similar explanation. Basketball point spreads are used secondarily by non-wagering sports viewers and media to forecast the margin of victory and help determine whether an upset has occurred (Camerer, 1989; Brown & Sauer, 1993). The formula for individual basketball referee z-scores based on point spreads (zPS) is as follows:

$$zPS = \mu_{\text{residual per ref}} \div \left(\sigma_{\text{residuals all games}} \div \sqrt{\text{games per ref}} \right) \quad (3)$$

where $\mu_{\text{residual per ref}}$ is the average “point spread” residual for each game a specific referee worked, $\sigma_{\text{residuals all games}}$ is the standard deviation of the “point spread” residuals of all games in the dataset and $\sqrt{\text{games per ref}}$ is the square root of the number of games officiated by each referee.

The Dallas Mavericks provide a good test case for this type of analysis. Team owner Mark Cuban has been fined by the NBA at least ten times since 2000 for no less than a total of \$US1.45 million (DuPree, 2006). The vast majority of Cuban’s fines have been attributable to his criticism of referees. On or around January 4, 2001, he said: “[t]he refs were pitiful tonight and I don’t care if I get fined” (Associated Press (AP), 2006). On or around January 8, 2002, Cuban declared: “[former NBA Director of Officials] Ed Rush might have been a great ref, but I wouldn’t hire him to manage a Dairy Queen. His interest is not in the integrity of the game or improving the officiating” (AP, 2006). On May 7, 2006, he blogged: “Refs miss calls. Its [sic] part of the game. Better refs that approach the game objectively miss fewer calls” (Cuban, 2006). Individual referees and the National Basketball Referees Association—the referees’ exclusive labor union of which all referees are members—surely take note of such comments.

NBA referees have a checkered past, lending anecdotal support for the importance of conducting this research. In 1997, NBA official Joey Crawford and sixteen other referees were indicted for federal tax evasion in connection with a frequent-flier mileage scam (AP, 1998). The same Joey Crawford was suspended by the NBA in 2007 for “improper conduct” (NBA Media Ventures, 2007) involving an altercation with San Antonio Spurs star Tim Duncan. After the incident, apparently sensing the possibility of retribution from other officials, Duncan expressed hope that he and his team “get a fair shake” and avoid any “backlash” (AP, 2007) in future games. More so than any other incident, the much-publicized gambling scandal involving long-time NBA referee Tim Donaghy exposed the extent to which NBA officials are capable of nefarious conduct. A post-Donaghy internal review revealed that a majority of NBA referees were in violation of their employment contract’s ban on gambling, although Commissioner Stern said none of the violations were major (Sheridan, 2007). Pedowitz (2008) confirmed such findings, highlighting the fact that 52 of the 57 referees interviewed during his investigation acknowledged that they had engaged in some form of impermissible gambling while contemporaneously working as an NBA-employed official.

In the wake of Donaghy's venal conduct, the NBA made a number of moves to bolster its referee oversight. First, the NBA created the new position of senior vice president for referee operations and filled it with retired US Army General Ronald L. Johnson. Secondly, as detailed by Pedowitz (2008, p. 44), the NBA developed a "system of recruiting, training, monitoring, managing, and developing" its referees. Thirdly, the NBA hired Steven Angel as Senior Vice President for League Operations and Officiating to "help coordinate wagering intelligence and game screening" (Pedowitz, 2008, pp. 113–114). All of these measures came several years after the NBA started using digital technology in its review of game film to monitor individual referees and evaluate specific calls (and non-calls) being made.

Upon a statistically significant finding of referee-specific adverse effect on team performance, retribution theory is the most likely explanation for any possible bias by individual NBA referees. Rawls (1955, p. 4) defined retribution as follows: "What we may call the retributive view is that punishment is justified on the grounds that wrongdoing merits punishment". The retribution-related work of Wittman (1974) and Posner (1980) is most relevant here. Wittman (1974, p. 233) elaborated on how retribution "predicts a considerable amount of behavior [and] explains certain cases of failure by other approaches to punishment". Along these lines, NBA referees could plausibly view the NBA-imposed fines on the multi-billionaire owner of the Dallas Mavericks as insufficient punishment for his critical remarks and seek additional punishment through "payback" calls (or non-calls) against the Dallas Mavericks. NBA referees would likely rationalize their lack of objectivity as justified retribution for Cuban's perceived wrongdoing.

In examining retribution and other related theories of punishment, Posner (1980) made two important points that are germane here. First, he discussed the collective nature of retribution, which is particularly applicable to the context of an employee who is also a member of a labor union, as is the case with NBA referees. An attack, verbal or otherwise, by Mark Cuban directed towards one individual referee may be construed by other referees as an insult at them and/or the entire union membership. Secondly, in the course of showing how retribution can be explained economically, Posner (1980, p. 92) opined that "people may obtain utility from inflicting costs on people who have injured them". In this way, an NBA referee "injured" by Cuban's comments who is engaged in retribution against the Dallas Mavericks may glean an intangible psychic benefit from officiating-related woes suffered by the team as a result of his or her bias.

Retribution is not a foreign concept to NBA referees. During every game played on February 27, 2004, in protest of fellow NBA referee Michael Henderson's three game suspension for inadvertently blowing his whistle at a critical time during a game earlier in the month, all of the NBA referees turned their uniforms inside out and wrote Henderson's number 62 on the back. The visible and very public response to the NBA's suspension of Henderson for what was an unintentional error in judgement drew a sharp rebuke from Commissioner Stern. The following day, he said: "Last night's

display was woefully inconsistent with the professionalism with which NBA officials normally conduct themselves” (SportsTicker, 2004). The findings in the Pedowitz Report also support the theoretical foundation for this analytical model. Pedowitz (2008, p. 56) found that a “substantial number of [NBA] team representatives believes that referees make calls, on occasion, based on personal bias”. Pedowitz (2008, p. 59) also acknowledged that “the perception that bias exists undermines confidence in the integrity of the game”.

Data and Results

Data were gathered via box scores for the Dallas Mavericks over the course of seven seasons starting in 2001–2002 and ending in 2007–2008.² During this time period, the Mavericks played 654 games. Of the total, 574 were regular season games and 80 were playoff games. The actual score of each game was recorded and the names of all three game officials were noted. A total of 77 different NBA referees officiated games during this seven-year period. Year-end Sagarin ratings from the *USA Today* were compiled for both the Dallas Mavericks and their opponents each season, with the appropriate home/away adjustment considered in calculating the predicted margin of victory. In addition, point spreads from Covers.com were included for each game. Both the Sagarin-derived prediction and point spread were then compared to the actual score differential. Table 1 displays summary descriptive statistics for: (1) Sagarin’s predicted score versus the actual score of all 654 games (SAG_{all}) and for only the 80 playoff games ($SAG_{playoffs}$) and (2) the point spread versus the actual score of all 654 games (PS_{all}) and just the 80 playoff games ($PS_{playoffs}$).

Multiple z-scores were calculated for each referee who officiated at a Dallas Mavericks game during the seven-year period. In an effort to make the results as robust as possible, the z-scores specific to each referee were determined in three different ways.³ For the subset of referees who officiated playoff games contested by the Mavericks, additional z-scores on each of the same three bases were calculated. Table 2 sets forth the distribution of z-scores based on Sagarin’s predicted margin of victory ($zSAG_{all}$), Sagarin’s predicted win–loss outcome (zWL_{all}) and point spreads (zPS_{all}) for all games. Table 3 summarizes the $zSAG_{all}$, zWL_{all} and zPS_{all} of each referee for all games during the time period, with statistically significant negative z-scores

Table 1. Sagarin predicted score/point spread vs. actual score (all games and playoffs only)

	SAG_{all}	$SAG_{playoffs}$	PS_{all}	$PS_{playoffs}$
Mean	0.03	−1.01	0.24	−0.83
Standard error	0.46	1.39	0.46	1.44
Standard deviation	11.69	12.47	11.81	12.86
Kurtosis	0.23	0.03	0.12	−0.08
Skewness	−0.09	0.22	−0.08	0.19

Table 2. Distribution of SAG, WL and PS z-scores (all games and playoffs only)

	$zSAG_{all}$	zWL_{all}	zPS_{all}	$zSAG_{playoffs}$	$zWL_{playoffs}$	$zPS_{playoffs}$
Mean	-0.03	0.34	0.06	-0.18	-0.24	-0.15
Standard error	0.12	0.12	0.12	0.18	0.19	0.18
Standard deviation	1.03	1.06	1.06	1.10	1.16	1.12
Kurtosis	1.34	-0.65	0.83	0.47	0.20	0.00
Skewness	-0.65	-0.31	-0.44	0.20	-0.35	0.08

at the 5% level noted.⁴ Like any standardized scores, there are a small number of outliers that are significant. However, when considering all games, no official had a $zSAG_{all}$, zWL_{all} and zPS_{all} of less than negative two (-2). Only two referees had statistically significant z-scores on two out of the three measurements.

From 2001–2002 to 2007–2008, the Dallas Mavericks made the playoffs as one of the eight best teams in the NBA's Western Conference every year. During this time, they played in 80 playoff games, winning 38. A total of 37 referees worked these games. Table 2 sets forth descriptive statistics for $zSAG_{playoffs}$, $zWL_{playoffs}$ and $zPS_{playoffs}$ distributions in connection with playoff games only. Table 4 lists the $zSAG_{playoffs}$, $zWL_{playoffs}$ and $zPS_{playoffs}$ for each official who worked at least one Dallas playoff game during the time period, with significant z-scores flagged.⁵ The playoffs-only analysis revealed one referee who had statistically significant z-scores of less than negative two (-2) on all three measurements. Specifically, this referee's $zSAG_{playoffs}$ was -2.27 , $zWL_{playoffs}$ was -3.35 and $zPS_{playoffs}$ was -2.15 . Another referee had a $zSAG_{playoffs}$ of -2.16 , $zWL_{playoffs}$ of -1.95 and $zPS_{playoffs}$ of -2.32 .

Conclusion

When considering all three z-score measurements ($zSAG$, zWL and zPS) and all regular season and playoff games over the course of seven seasons, we did not find any NBA referee having a significant adverse effect on the Dallas Mavericks' team performance. However, when these data were bifurcated to solely analyze the more meaningful playoff games, one official's z-scores were significant on all three measurements. As such, it can be concluded that this official had a significant negative impact on the Mavericks' performance during the playoffs. Retribution is a likely explanation for any possible bias exhibited by such referee. These findings, however, must be reconciled with the understanding that indirect methods, such as those relied on here, are not conclusive. Bernhardt and Heston (2009) outlined the danger of such methods. In the absence of collaboration in the form of direct evidence, the *prima facie* showing could be rebutted by an innocuous explanation. The analysis here merely lends itself to the formation of a rebuttable presumption. Without more, such analysis is uncorroborated. A subsequent investigation that includes personal interviews, polygraph tests or the like would be necessary to conclusively ascertain whether any insidious conduct

Table 3. Referee z-scores ($zSAG_{all}$, zWL_{all} , zPS_{all}) for all games (regular season and playoff games)

Referee	$zSAG_{all}$	zWL_{all}	zPS_{all}
1	0.34	2.37	0.12
2	-0.78	-0.55	-0.73
3	0.06	-0.83	-0.25
4	-1.33	-0.40	-1.09
5	-0.97	-0.89	-1.18
6	1.95	1.30	1.87
7	0.98	1.33	0.31
8	0.92	1.71	0.64
9	-1.18	-1.85	-1.88
10	0.66	1.00	-0.16
11	-0.82	-1.46	-1.19
12	-0.65	0.05	-0.43
13	0.36	0.54	0.44
14	-0.44	1.20	-0.64
15	-0.34	-0.17	-0.26
16	0.13	0.31	0.28
17	0.41	0.48	0.42
18	-0.20	1.10	-0.25
19	0.40	0.97	0.27
20	1.43	1.77	1.66
21	0.39	1.13	-0.04
22	-1.33	-0.17	-1.77
23	0.40	1.03	0.01
24	-0.36	1.05	-0.37
25	0.36	-0.43	0.79
26	0.02	-0.12	0.37
27	-0.85	-0.65	-0.85
28	2.05	2.68	1.86
29	1.58	1.38	2.31
30	-0.61	0.68	-0.73
31	1.28	1.24	1.58
32	0.59	0.26	1.10
33	0.36	1.56	0.28
34	0.29	0.68	0.42
35	-0.24	0.27	-0.47
36	-2.56 ^a	-1.55	-2.81 ^a
37	0.42	1.07	0.80
38	0.54	1.49	0.49
39	-0.89	-0.82	-0.82
40	-0.85	-1.09	-0.79
41	1.14	1.82	1.40
42	1.59	0.87	1.35
43	-0.50	0.58	-0.39
44	0.88	1.29	1.24
45	0.27	0.19	0.30
46	-1.66	-1.91	-1.16
47	0.03	0.56	-0.24

Table 3 (Continued)

Referee	$zSAG_{all}$	zWL_{all}	zPS_{all}
48	-0.66	-0.65	-1.00
49	-0.15	-0.34	-0.28
50	-0.06	0.59	0.34
51	0.37	-0.06	0.02
52	-0.95	-0.53	-0.89
53	-0.23	0.72	-0.21
54	1.56	1.75	1.73
55	-0.37	-0.78	-0.66
56	-0.10	1.15	-0.54
57	-2.30 ^a	-1.43	-2.44 ^a
58	0.06	1.11	0.11
59	0.31	0.81	0.09
60	-0.43	-1.22	-0.24
61	1.65	1.36	1.78
62	-1.80	-1.43	-1.58
63	-1.56	-1.34	-1.42
64	-0.07	1.29	-0.07
65	0.77	0.25	0.42
66	0.44	0.59	0.45
67	0.92	0.80	0.60
68	0.38	-0.12	0.38
69	-3.68 ^a	-1.52	-3.29 ^a
70	-0.88	1.02	-0.91
71	0.67	0.79	0.97
72	0.34	0.14	0.51
73	-0.63	0.29	-0.46
74	-0.74	0.02	-0.38
75	-0.44	-1.33	-0.68
76	1.13	1.48	1.27
77	0.89	1.60	0.44

^a $p < 0.05$.

occurred. Nevertheless, with employee wrongdoers in all industries looking to conceal their activity in order to minimize detection by their employer and avoid getting fired, sued or charged with a crime (Duggan and Levitt, 2002; Bernhardt and Heston, 2009), it is plausible to consider that if an NBA referee sought retribution against Mark Cuban and/or the Dallas Mavericks, he or she might aim to “dilute” his or her malfeasance during the critical playoff games with more favorable officiating during the regular season, making it harder to detect.

Sport governing bodies such as the NBA should avail themselves of every reasonable, feasible and permissible resource in the course of monitoring the performance and objectivity of their employee referees. Commissioner Stern seems inclined to do so in the case of the NBA, as he has vowed to implement the recommendations contained in the Pedowitz Report.⁶

Table 4. Referee z-scores ($zSAG_{\text{playoffs}}$, zWL_{playoffs} , zPS_{playoffs}) for playoff games only

Referee	$zSAG_{\text{playoffs}}$	zWL_{playoffs}	zPS_{playoffs}
1	0.32	0.61	0.37
2	-1.11	-1.20	-1.12
3	-1.11	-2.06 ^a	-1.29
4	0.89	0.31	1.00
5	0.10	-0.71	0.04
6	0.99	0.52	0.96
7	-1.42	-0.67	-1.58
8	-2.27 ^a	-3.35 ^a	-2.15 ^a
9	2.84	1.53	2.66
10	-0.76	-1.15	-0.52
11	0.13	0.79	0.07
12	0.51	0.69	0.46
13	1.15	0.84	1.36
14	-0.08	-0.74	-0.20
15	1.44	2.24	1.31
16	0.45	-0.12	0.96
17	-0.82	-0.77	-0.90
18	0.54	0.43	0.63
19	-0.32	0.28	-0.68
20	0.52	0.68	0.66
21	-1.59	-1.80	-1.50
22	-0.46	0.35	-0.48
23	-1.20	-1.51	-1.09
24	0.47	1.12	0.58
25	-2.16 ^a	-1.95	-2.32 ^a
26	-0.50	-1.15	-0.58
27	-0.28	-0.28	-0.18
28	-2.12 ^a	-1.16	-2.12 ^a
29	-0.24	-0.72	-0.17
30	0.45	0.66	0.50
31	-0.30	-0.43	-0.38
32	-1.22	-0.76	-1.19
33	1.41	0.74	1.41
34	0.05	0.31	0.22
35	0.12	-0.22	-0.22
36	-1.16	-1.42	-1.06
37	0.11	1.21	0.33

^a $p < 0.05$.

Thomson (2008, p. 99) posited that “[NBA] fans have already grown too cynical about the refereeing to be satisfied with anything short of substantive change”. Ayres (2007, pp. 4–5) advocated that “the NBA should release its refereeing data and let forensic econometricians search for impropriety . . . Stern should be willing to put his refs to the test”. While the NBA is unlikely to acquiesce to such a transparent analysis, given that such data represent confidential employee records, the model proposed in this paper could be

used by certain basketball leagues as an additional non-invasive tool in the evaluation of individual referees and their interactions with specific teams, owners, coaches or players. In fact, Pedowitz (2008, p. 114) acknowledged that the NBA is currently developing statistical analyses to “help identify patterns consistent with referee bias for/against certain players or teams”.

There are a number of limitations to this paper. First, data were not analyzed for any time period prior to Mark Cuban’s ownership of the Dallas Mavericks. Such an analysis could potentially serve as a “control group” and make the results here more robust by shedding light on whether any referee-based adverse effect on team performance or bias exhibited by referees is more specific to Mark Cuban or applicable to the Dallas Mavericks franchise generally. Secondly, no analysis was extended to other NBA teams during the same time period. Such an exercise would reveal whether certain referees affect the performance of other teams in similar ways. Thirdly, no attempt was made to control for changes in the Dallas Mavericks’ coaching staff or team composition, as the presence or absence of certain personnel affiliated with the team could be relevant. Fourthly, as specifically alluded to in the title of this paper, the analysis here is merely the “starting point”. In order to corroborate the findings derived from the calculation of referee-specific z-scores, a regression analysis to explain team performance as a function of certain variables and a dummy variable for referees would be necessary.

This research line could be extended in several other directions. Specific calls of individual referees could be analyzed at the game-level to glean further insight into how such officials affect team performance. This model could also be altered to measure three-person referee squads in addition to individual referees. Further, while the purpose of this paper is to measure adverse impact on team performance and possible bias being exhibited by NBA referees (hence the exclusive focus on statistically significant negative z-scores), a related investigation could be undertaken to ascertain whether any NBA referees are biased in favor of certain teams, owners, coaches or players and positively impact team performance. In this way, non-conclusive evidence of favoritism would be inferred from statistically significant positive z-scores for each respective basketball referee. As was done by Larson, Price and Wolfers (2008) in the context of racial bias by NBA referees, this research could also be extended to determine whether a profitable betting strategy could be developed based on our findings. Finally, this model could be adapted for use in sports other than basketball, where the performance of officials must be constantly reviewed as a gauge of competency and impartiality.

Notes

1. NBA Commissioner David Stern upon the release of the Pedowitz Report on October 2, 2008 (NBA Communications Group, 2008).
2. Mark Cuban became majority owner of the Dallas Mavericks in January of 2000.
3. See equations (1)–(3) previously stated.

4. Given the collective nature and macro-level analytical focus of this paper, Table 3 lists individual referees by non-identifying number, not name.
5. Like Table 3, Table 4 lists individual referees by non-identifying number, not name.
6. In AP (2007), Stern stated: "There is nothing as important as the integrity of our game and the covenant we have with our fans. In order to preserve their trust, we will make every effort to ensure that our processes and procedures are the best they can be".

References

- Associated Press (1998). Referee resigns; Crawford pleads guilty to tax fraud, *CNN.com*, June 25. Retrieved May 11, 2009 from http://robots.cnn.com/basketball/nba/news/1998/06/24/referee_taxes/
- Associated Press (2006). Timeline of Mark Cuban fines, *ESPN.com*, June 20. Retrieved May 11, 2009 from <http://sports.espn.go.com/espn/wire?section=nba&id=2492950>
- Associated Press (2007). Spurs, Duncan worry about ref backlash, *NBCsports.com*, April 19. Retrieved May 11, 2009 from <http://nbcsports.msnbc.com/id/18124750/>
- Ayres, I. (2007). Give Freakonomics a chance. *Economists' Voice*, November, pp. 1–5. Retrieved May 11, 2009 from <http://www.bepress.com/cgi/viewcontent.cgi?article=1299&context=ev>
- Bernhardt, D., & Heston, S. (2009). Point shaving in college basketball: A cautionary tale for forensic economics. *Economic Inquiry*, forthcoming.
- Berri, D.J., & Jewell, T. (2004). Wage inequality and firm performance: Examining a natural experiment from professional basketball. *Atlantic Economic Journal*, 32(2), 130–139.
- Berri, D.J., Leeds, M.A., Leeds, E.M., & Mondello, M. (2009). The role of managers in team performance. *International Journal of Sport Finance*, 4(2), 75–93.
- Brown, W.O., & Sauer, R.D. (1993). Fundamentals or noise? Evidence from the professional basketball betting market. *The Journal of Finance*, 48(4), 1193–1209.
- Camerer, C.F. (1989). Does the basketball market believe in the 'hot hand'? *American Economic Review*, 79(5), 1257–1261.
- Carmichael, F., Thomas, D., & Ward, R. (2000). Team performance: The case of English Premiership football. *Managerial and Decision Economics*, 21, 31–45.
- Cuban, M. (2006). How to improve NBA playoff officiating, *blog maverick*, May 7. Retrieved May 11, 2009 from <http://www.blogmaverick.com/2006/05/07/how-to-improve-nba-playoff-officiating/>
- Dare, W., & MacDonald, S. (1996). A generalized model for testing the home and favorite team advantage in point spread markets. *Journal of Financial Economics*, 40(1996), 295–318.
- Duggan, M., & Levitt, S.D. (2002). Winning isn't everything: Corruption in sumo wrestling. *American Economic Review*, 92(5), 1594–1605.
- DuPree, D. (2006). NBA fines Cuban \$250,000 for behavior after game 5, *USAToday.com*. Retrieved May 11, 2009 from http://www.usatoday.com/sports/basketball/playoffs/2006-06-20-cuban-fine_x.htm
- Espitia-Escuer, M., & Garcia-Cebrian, L.I. (2006). Performance in sports teams: Results and potential in the professional soccer league in Spain. *Management Decision*, 44(8), 1020–1030.
- Hadley, L., Poitras, M., Ruggiero, J., & Knowles, S. (2000). Performance evaluation of National Football League teams. *Managerial and Decision Economics*, 21, 63–70.
- Harris, M.B. (1998). *Basic statistics for behavioral science research*. Upper Saddle River, NJ: Allyn & Bacon.
- Harville, D. (1977). The use of linear-model methodology to rate high school or college football teams. *Journal of the American Statistical Association*, 72(358), 320–338.
- Harville, D. (1980). Predictions for National Football League games with linear model methodology. *Journal of the American Statistical Association*, 75(371), 516–524.
- Kahn, L.M. (1993). Managerial quality, team success, and individual player performance in Major League Baseball. *Industrial and Labor Relations Review*, 46(3), 531–547.

- Larson, T., Price, J., & Wolfers, J. (2008). Racial bias in the NBA: Implications in betting markets. *Journal of Quantitative Analysis in Sports*, 4(2), Art. 7. Retrieved May 11, 2009 from <http://www.bepress.com/cgi/viewcontent.cgi?article=1112&context=jqas>
- NBA Communications Group (2008). Pedowitz report recommends significant changes, *NBA.com*, October 2. Retrieved May 11, 2009 from <http://www.nba.com/2008/news/10/02/nba.release/index.html>
- NBA Media Ventures (2007). NBA referee Joey Crawford suspended, *NBA.com*, April 17. Retrieved May 11, 2009 from http://www.nba.com/news/crawford_070417.html
- Oorlog, D.R. (1995). Serial correlation in the wagering market for professional basketball. *Quarterly Journal of Business and Economics*, 34(2), 96–109.
- Pedowitz, L.B. (2008). Report to the board of governors of the National Basketball Association, *NBA.com*, October 1. Retrieved May 11, 2009 from <http://www.nba.com/media/PedowitzReport.pdf>
- Posner, R.A. (1980). Retribution and related concepts of punishment. *Journal of Legal Studies*, 9(1), 71–92.
- Price, J., & Wolfers, J. (2007). *Racial discrimination among NBA referees*. Cambridge, MA: National Bureau of Economic Research Working Paper 13206.
- Rawls, J. (1955). Two concepts of rules. *Philosophical Review*, 64, 3.
- Schnytzer, A., & Weinberg, G. (2004). Is the NBA betting market efficient? In G.T. Papanikos (Ed.), *The economics and management of mega athletic events: Olympic games, professional sports, and other essays* (pp. 1–15). Athens, Greece: ATINER.
- Scully, G.W. (1974). Pay and performance in Major League Baseball. *American Economic Review*, 64, 915–930.
- Sheridan, C. (2007). NBA to revamp ref gambling rules: Jackson, Nunn see roles reduced, *ESPN.com*, October 26. Retrieved May 11, 2009 from <http://sports.espn.go.com/nba/news/story?id=3079309>
- SportsTicker (2004). Stern upset over protest by NBA refs, *NBCSports.com*, February 29. Retrieved May 11, 2009 from <http://nbcsports.msnbc.com/id/4402659/>
- Thomson, I. (2008). 5 minute guide. *Sports Illustrated*, October 27, pp. 98–101.
- Windhorst, B. (2009). NBA insider: Going way behind the box score. *Cleveland Plain Dealer*, March 8. Retrieved May 22, 2009 from http://www.cleveland.com/cavs/index.ssf/2009/03/nba_insider_going_way_beyond_r.html
- Winston, W.L. (2007). *Excel 2007: Data analysis and business modeling*. Redmond, WA: Microsoft Press.
- Wittman, D. (1974). Punishment as retribution. *Theory and Decision*, 4(3/4), 209–237.