Point Shaving in NCAA Basketball: Corrupt Behavior or Statistical Artifact?

by

George Diemer

Mike Leeds

Department of Economics

DETU Working Paper 10-09

August 2010

1301 Cecil B. Moore Avenue, Philadelphia, PA 19122

http://www.temple.edu/cla/economics/

Point Shaving in NCAA Basketball: Corrupt Behavior or Statistical Artifact?

By George Diemer and Mike Leeds

Introduction

Gambling on sporting events has probably existed as long as sports themselves. The founding myth of the ancient Olympic Games surrounds a chariot race that the Greek hero Pelops rigs so as to win the race and marry the king's daughter. Now, as then, gamblers have an incentive to insure that they win their bets on athletic events by somehow fixing the outcome.

Gambling and tampering have accompanied the development of team sports in America. According to Seymour (1961: 53), the Mutual Club of New York intentionally lost a game to the Eckford Club of Brooklyn in 1865. This "fix" occurred four years before the first overtly professional team suited up, eleven years before the first organized professional league in North America, and over a half-century before the "Black Sox" scandal, in which Chicago players "threw" the 1919 World Series.

We look at one particular form of fixing that has received considerable attention in both the popular press (See, for example, McCarthy, 2007) and economic literature (See Wolfers, 2006; Borghesi (2008); Johnson, 2009; Diemer (2009B); and Bernhardt and Heston, 2010): point shaving in college basketball. In particular, we use new techniques and datasets to confirm Wolfers' claim that evidence suggests that point shaving exists for regular season NCAA college basketball games. We then test whether evidence of point shaving declines as the incentives to shave decrease. We do this by looking for evidence of point shaving during post-season tournament play. Our results provide further support for the hypothesis that point shaving takes place during the regular season.

In the next section of the paper, we explain point shaving and review the literature on it. In section three, we present our empirical model and describe our data set. In section four, we use the distributions of the outcomes of regular season games relative to the point spread to reject the null hypothesis of no point shaving. We fail to reject the null hypothesis of no point shaving in tournament play. Section five concludes.

II. Literature Review

Point Shaving

Point shaving typically occurs when one or more participants ensure that his team wins by less than the point spread.[1] The point spread is the number of points by which gambling houses expect the favored team to beat the underdog. Point spreads are a popular form of gambling because they attract bets on games between two unevenly matched teams. A team can lose the game badly, but it can still garner support by gamblers if they can bet that the team will lose by less than the bookmakers predict.

College basketball is particularly susceptible to point shaving for several reasons. First, because of the monopsony power exercised by the NCAA, college athletes are not compensated for their efforts beyond scholarships and highly limited aid. In the case of elite athletes in the marquee sports, this compensation can be a small fraction of the revenue that the players generate for their colleges. (See Brown, 1994.) In addition, college basketball players are often less prepared for college than other students and

---

[1] Point shaving is thus not restricted to players, as coaches, trainers, and referees can all affect the game's outcome.

graduate at lower rates than their classmates. (See Leeds and von Allmen, 2010.) This, in turn, reduces an athlete's attachment to his university and the repercussions of point shaving for these athletes.[2]

The nature of basketball lends itself to point shaving. Basketball is a much higher-scoring game than sports like soccer or baseball. As a result, the final point differential is generally much larger in basketball. The wider differential increases the opportunity for allowing the other team to score or to failing to score oneself without greatly increasing the chance that one will lose the game.[3] In addition, unlike football, relatively few points can be scored in any one possession, which limits the ability of a team to make up a large deficit very quickly. Finally, the large number of regular season games reduces the amount of scrutiny any one game receives.

Point Shaving and the Efficient Markets Hypothesis

The economic literature treats gambling as it would any financial market in that it assumes that the Efficient Markets Hypothesis (EMH) applies. The EMH proposes that the price of any tradable commodity incorporates all available information. In the gambling market the tradable commodity is the wager itself, and the price is the point spread. Therefore, for the EMH to apply, the point spread has to incorporate all available information about the game. For gambling markets to be efficient, the difference between game outcomes and point spreads has to have a median of zero over a large enough sample size (Wolfers and Zitzewitz, 2004). For more information on the point

[2] Universities can cancel scholarships or expel the student athlete.
[3] Apparently, point shaving almost always involves the favorite's winning by less than the point spread and almost never involves the underdog's losing by more than the point spread. Wolfers (2006) attributes this to the fact that the former form of point shaving does not interfere as much with a player's competitive instincts.

spread see Diemer (2009A), for more information on market efficiency see Vaughan Williams (2005).

Wolfers (2006) tests for the existence of point shaving in college basketball from the 1989-1990 season to the 2004-2005 season for a sample of 44,120 games. Regressing the actual winning margin on the point spread yields the first column of Table 1:

This result shows that the point spread is a good overall predictor of the game outcome, as the EMH predicts. Next, Wolfers examines the outcomes of the subset of games that have point spreads of 12 points or more. If gambling markets are efficient and the probability density function (PDF) of game outcomes is normally distributed, the distribution of winning margins should be symmetric and peak at the point spread. Normally distributed game outcomes – and the resulting symmetry absent point shaving – are a critical assumption in Wolfers and most subsequent research. Much of the previous literature bases its conclusions on tests for distributional symmetry.

To test the null of no point shaving in college basketball, Wolfers tests whether the distribution of game outcomes is symmetric:

$$p(0 < \text{Winning margin} < \text{Spread}) = p(\text{Spread} < \text{Winning margin} < 2 * \text{Spread}) \quad (1)$$

Wolfers finds that a disproportionate percentage of heavy favorites wins the game but fails to cover the point spread (the left side of the equation is significantly larger than the right side of the equation). As a result, he concludes that “point shaving led roughly 3 percent of strong favorites who would have covered the spread not to cover (but still win)" (p. 282).

Bourghesi (2008) challenges Wolfers' interpretation. Bourghesi claims that economic agents in professional leagues have little incentive to shave points because they are so closely watched. He assumes symmetry, tests professional leagues as Wolfers had, and finds similar results. However, Bourghesi attributes his findings to the sports books' pursuit of profit: "sports books intentionally inflate the lines of games in which a particularly strong team plays against a particularly weak team. This practice, called line shading, potentially maximizes sports book revenues." (Borghesi, 2008: 1068)

Bernhardt and Heston (2010) derive their own point spreads for games that have no line posted. They then compared the game outcome relative to this derived point spread. They find that the "distinct asymmetric patterns … are driven by a common desire to maximize the probability of winning" by the players. (Bernhardt and Heston: 15). The conclusion is that strategic game play is the cause of the statistical anomaly.

Johnson (2009) points out that Wolfers' analyzing games with heavy favorites introduces a regression effect that biases the results. The regression effect occurs when bettors' actions drive the point spreads too high. A strong regression effect causes heavy underdogs to cover more often. Forcing the game to have a winner and a loser (basketball games cannot end in a tie), distorts the distribution of outcomes for slight favorites. For example if a team is favored by two points it cannot lose the wager by two points. This disturbance might cause researchers to falsely reject the hypothesis of no point shaving.

Diemer (2009B) confirms Wolfers' findings by examining gambling on National Football League (NFL) games. He constructs and tests non-parametric PDFs of game outcomes relative to the point spreads, discarding the assumption that winning margins are distributed normally (and symmetrically). Diemer uses a bootstrap test to determine whether the distribution of winning margins for slight favorites is the same as the distribution of winning margins for heavy favorites. If the gambling market is efficient, the size of the spread should not alter its accuracy as a predictor of the final outcome, so the distributions of the outcomes for heavy and slight favorites should be equal. This equality, in turn, is a sufficient condition of accepting the null of no point shaving. Diemer uses this test of equality to test for evidence of point shaving in NFL regular season games from 1993 to 2007. Diemer rejects the null hypothesis of equality. Instead, he finds that, "While the point spread is a function of perceived outcome, the bimodal distribution [of the heavy favorites] suggests the outcome is a function of, in part, the point spread itself. …As the point spread increases, so does the incentive to shave points, which leads to increasingly skewed densities [of the heavy favorites]." (pp. 22-23)

III. Empirical Model and Data

We collected data from www.goldsheet.com from the 1995-1996 season through 2008-2009 for NCAA basketball games. We deleted "pick 'em" games in which two teams are presumed to be evenly matched, and bettors must pick the winner (an effective point spread of zero). Analyzing a game with no favored team is counter to the objective of this research, so these games were deleted. In addition, point spreads are posted only for games that are likely to attract betting interest. This leaves a sample of 31,793 regular season and 3,371 playoff games.

As noted above, we compare game outcomes to the point spread. We do so by defining Net Favored Points (NFP) as:

NFP = Favored team's points scored – Underdog's points scored + Point Spread according to the favored team (3)

If the favored team covers the point spread, NFP > 0; if it fails to cover the point spread NFP <0. If the point spread accurately predicted the outcome, NFP = 0. According to the Efficient Markets Hypothesis, the distribution of the NFP should peak where NFP = 0.

Efficient Markets and the Probability Density Functions

We first conduct a parametric test of the aggregate dataset to test the EMH by constructing kernel probability density functions (PDF's) of the NFP, using a normal optimal smoothing parameter. Figure 1 displays the PDF of all 35,164 NCAA game outcomes (both regular season and post-season tournament games). The band around the distribution shows the 95-percent confidence interval around the distribution.[4] When the distribution falls within the shaded area of the reference band the distribution is normal. As shown in Figure 1 the distribution falls within the reference band and is centered in the neighborhood of zero. This normal distribution with a peak at zero indicates that the

[4] The shaded area barely visible in this instance.

point spread is the best aggregate predictor of the game outcomes, as we would expect in an efficient gambling market.

Testing for Point Shaving in Regular Season NCAA Games

As noted above, point shaving is more likely to occur at higher point spreads. As the spread increases, it becomes easier for players to satisfy their secondary objective (winning the wager) without compromising their presumed primary objective (winning the game). Like Diemer (2009B), we conduct a non-parametric test for equality of two distributions: heavy favorites (f(NFP)) and slight favorites (g(NFP)).[5] The formal hypothesis tests are:

H_o: f(NFP) = g(NFP), for all NFP

H_1: f(NFP) $\neq$ g(NFP), for some NFP

A bootstrap test of equality generates a p-value as a global indicator of equality. We also provide a reference band around each NFP to illustrate the hypothesis test. The band is two standard errors wide at any NFP (which is measured along the horizontal axis). If the densities fall outside the reference band, that portion of the distribution is a likely reason for rejecting H_0. The normal optimal smoothing parameter is the standard, as it minimizes the risk of falsely rejecting the null of equality.[6] We performed this test in two stages, first for 31,793 regular season games and then for 3,371 playoff games.[7]

---

[5] Eliminating the parametric assumption of the dataset eliminates an added constraint. Instead of testing for parametric characteristics of subsets of a dataset we test the equality of two subsets.
[6] A Sheather-Jones plug-in bandwidth was also run, which resulted in similar results.
[7] Page constraints limit the findings presented in this research. All point spread thresholds are available upon request of one of the authors.

IV. Results

The results of our tests appear in Figures 2, 3, 4, 6, and 7. Figure 2 divides regular season games into winning margins for heavy favorites (dashed lines) of 19 points or more and slighter favorites (solid line) of 18.5 points or less. The global test for equality produces a p-value of 0.00, indicating that we can reject the null hypothesis that the two distributions are the same at the one percent level, which is one of the many necessary conditions for rejection of the overall null of no point shaving presented in this paper. We next examine why the global test failed.

The peak of the heavy favorites' PDF is at NFP = -4.3. This outcome is highly suspicious because it means that the heavy favorites win the game (they would lose if NFP < -18.5) yet are most likely to fail to cover the spread because NFP < 0. In addition the heavy favorites' distribution is bimodal, with peaks at NFP = -4.3 and 5.6 the valley at NFP = 2.3.

The bimodal PDF suggests that the point spread becomes less precise as it increases (a feature supported by our test for heteroskedasticity below). This finding is consistent with Diemer's (2009B) finding a bimodal distribution of outcomes in the NFL.

The skewness of the PDF also supports the claim that rejecting the null hypothesis of identical distributions results from an innocent statistical anomaly such as "strategic efforts," (Berhardt and Heston 2010) "line shading," (Borghesi 2008) or "regression to the mean" (Johnson 2009). However, if the causes of this statistical anomaly are as innocent as set forth above, the PDF would be symmetric even if it were not centered on zero. However, as shown in Figure 2, the heavy favorites' density (the dashed line) is far from symmetric. In addition, Figure 2 shows that the PDFs fall outside the reference band when 24.3>NFP>15.8. This supports the finding that "heavily favored teams appear to be involved in too many blowouts" (Wolfers, 2006: p.282). In other words, the heavy favorite either wins by far more than the point spread or fails to cover (but still wins the games).

Figure 3 displays the results at the 11.5-point threshold. The p-value is again 0.00, meaning we again reject H_0. This time, each PDF peak is located outside the reference band at NFP= -1 (for the heavy favorites), and NFP = 2.5 for the slighter favorites. The heavy favorites are again involved in too many blowouts at 27.8>NFP >19. Finally, the PDFs (especially the heavy favorites, as denoted by the dashed lines) become increasingly skewed as the points spread increases. The evidence of point shaving thus becomes stronger as the incentive to shave points (i.e., the point spread) increases.

Another reoccurring concern among those who doubt that point shaving exists is the tie game constraint. In this case, the lack of tie games skews the PDF for games with relatively slight favorites. Figure 4, which shows teams favored by 3.5 or more points and teams favored by 3 points or less, highlights this fact, resulting in a p-value of 0. For example, games with 3-point favorites cannot have outcomes for which NFP = -3. The relatively smaller favorite (the solid line) is bimodal due this the tie game constraint. The peaks are NFP = -3.9 and 2.5, and the valley is at NFP = -0.7, resulting in negative kurtosis for the smaller favorites. If we were able to correct for the tie game constraint in Figure 2 the density function of the slighter favorites would have a single, higher peak.

In other words, the games that create the solid, bimodal distribution shown in Figure 4 are also in the sample that produces the solid distributions in Figures 2 and 3, thus flattening those distributions. The tie game constraint thus understates the likelihood of point shaving.

Heteroskedasticity and Point Shaving in Regular Season Games

Our results thus far reject the hypothesis that the PDF of the NFP for heavy favorites is identical to the NFP for slight favorites in the regular season games. We also find that the likely cause of the inequality indicates point shaving. We further test the null of no point shaving by reexamining and reproducing Wolfers' basic equation for our sample of regular season NCAA games. As noted above, Wolfers finds that the actual point differential rises 1:1 with the point spread. He does not, however, test for the variance of this relationship. A constant variance in the regular season games would be consistent with the null hypothesis of no point shaving, but a variance that changes with the point spread would be further evidence that the point spread affects the margin of victory.

We perform this test by regressing the ex-post point differential on the point spread for all 31,793 regular season games in our sample. We then test the error term for heteroskedasticity. If we reject the null hypothesis of no heteroskedasticity, then we have stronger evidence of point shaving. If we cannot reject the null hypothesis, then the results contradict our previous finding.

The results of the regression appear in the second column of Table 1. While our results are not identical to Wolfers', we do replicate his key result, that the point differential rises roughly 1:1 with the point spread. The key finding in the first column of Table X, however, is not in the coefficients or R^2. Instead, it is in the result of the Breusch-Pagan test for heteroskedasticity. The test rejects the null hypothesis of a constant variance for regular season games at the 1-percent level of significance. Rejecting the null hypothesis shows that the distribution of the error structure varies with the point spread for regular season games. This finding supports our rejecting the null of no point shaving.

Tournament Play

Under the null hypothesis of no point shaving, the incentives to win a game are the same for tournament play as for regular season games.[8] All available information about the game outcome is incorporated in the point spread for both types of games. If an innocent statistical anomaly is the cause of the above statistical disturbances found by Wolfers (2006), the anomaly should hold true in both sets of games, and the PDFs of the NFP should be the same for them both.

Under the alternate hypothesis, point shaving should be less likely for tournament games. Because tournament games are more important than regular season games, the cost of losing a game is greater, and players will be more averse to behavior that increases the risk of losing. In addition, the greater attention that the public pays to tournament games increases the probability that point shaving will be detected.

Our sample consists of 3,371 tournament games. Tournaments are classified as any potentially season-ending games, such as NCAA Tournament, NIT Tournament, and

[8] If the incentive to win increases during tournament play it does so for both the favorite and the underdog uniformly.

Conference Tournaments. We again constructed and tested the PDFs and preformed the global test of equality for the PDFs of heavy favorites and slight favorites.

Results appear in Figures 5, 6, and 7. Figure 5 shows that the aggregate distribution is normally distributed, as the distribution falls inside the confidence band. Figures 6 and 7 test whether the PDFs of heavy and slight favorites are equal. As before, the point spread thresholds are 19 and 11.5 points. The resulting p-values –0.28 and 0.35 – do not allow us to reject H_0. The figures show that the distributions fall inside the reference band, so any fluctuations probably reflect white noise.

Heteroskedasticity and Point Shaving in Tournament Games

We next ran Wolfers' basic equation for NCAA tournament games. If the evidence of point shaving is a statistical artifact, then regular season games and tournament games should show similar results. Under the null of no point shaving, we would again expect to find heteroskedasticity for post-season games. If we cannot reject the hypothesis of a constant variance, then we have still more reason to believe that incentives to shave points again effect game outcomes. This time, as the incentives to shave points decreases in tournament play, statistical evidence of point shaving (heteroskedasticity) disappears.

The results of this regression appear in the third column of Table 1. As before, we find that the actual point differential rises 1:1 with the point spread. This time, however, the Breusch-Pagan fails to reject the null hypothesis of a constant variance. This provides further evidence that the distributions of the NFP is constant regardless of the point spread and that point shaving does not occur for tournament games.

V. Conclusion

Wolfers (2006) claim that point shaving is widespread in college basketball has drawn considerable attention in the literature, most of it attacking his findings. We perform several tests that support and build on Wolfers' results.

At first glance, our findings appear only partly consistent with Wolfers' (2006) claim that point shaving is common in college basketball. In particular, we find that his results hold for regular season games but not for post-season games. This result, however, actually provides greater support for Wolfers' basic claim, as we show that rational agents respond in the expected manner to incentives.

We show that point shaving is more likely to occur for regular season games that have larger point spreads. This satisfies the gamblers who want to fix the game while minimizing the change that the favored team will lose the game. Because post-season games receive much closer scrutiny than regular season games, they are likely to be much harder to fix undetected. Moreover, because post-season games are much more important than typical regular season games, we expect players to be more reluctant to participate in point shaving. Our results confirm this hypothesis. Our results thus reject the null hypothesis of no point shaving when the incentives to shave points are strong and fail to reject the null hypothesis when the incentives are weak.

While our discussion largely centers on players as the source of point shaving, we want to emphasize that players are only one possible source of game fixes. A coach who puts a less-skilled player in the game gets the same result as a player who does not play up to his ability. A referee who unjustifiably calls foul has he same impact as a player who unjustifiably commits a foul. Our results cannot identify the source of the point shaving other than to say it involves at least one person who can affect the game outcome. They only support the claim that point shaving exists in college basketball regular season games.

Figure 1: Empirical Probability Density Function: Aggregate Parametric Data

1995-2009

Probability Distribution: Favored Outcomes Relative to Point Spread
"x" axis measures the favored team's outcome relative to the point spread (NFP).
N = 35,164
Shaded area represents a normal distribution reference band

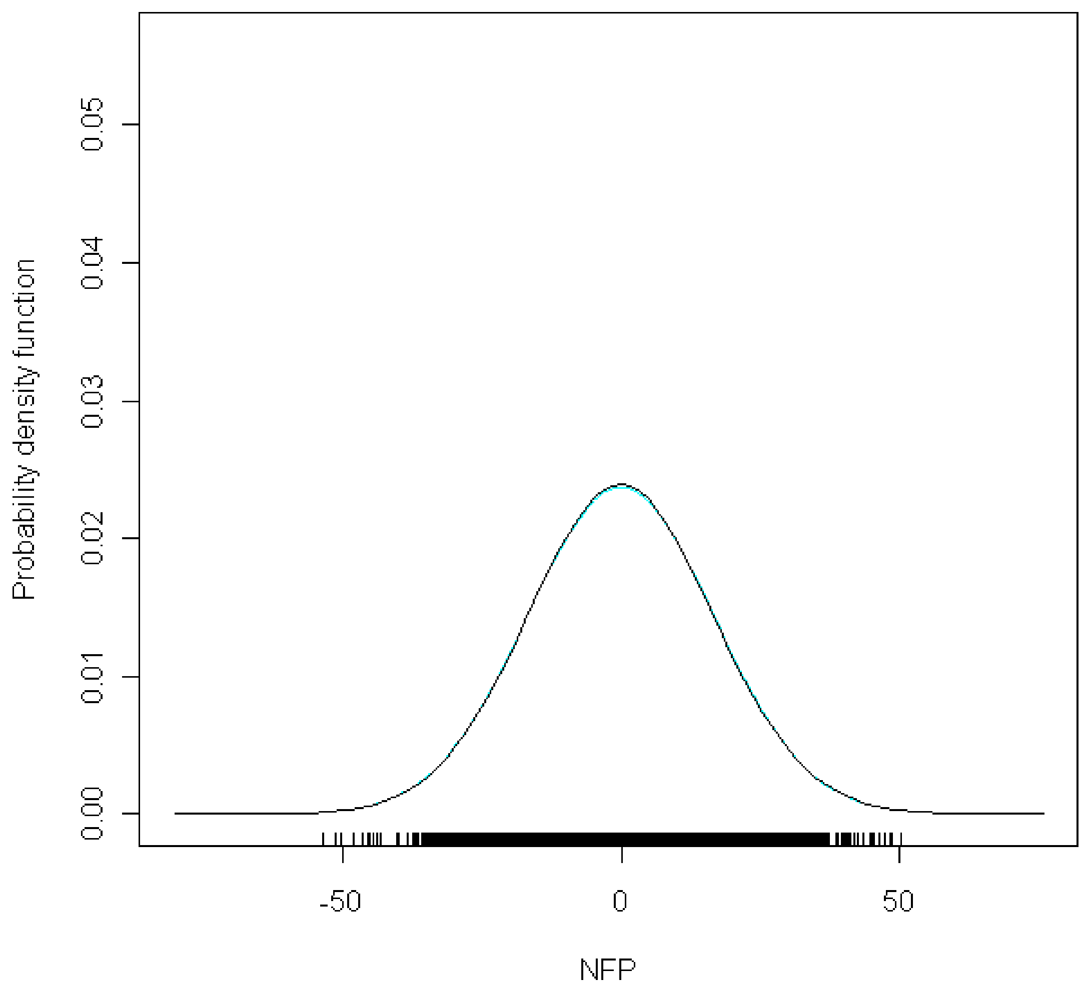

Figure 2: Regular Season Test of two distributions: 19 Point Favorites

Dashed line = Lower Bound favorites of 19 points or more; N = 2,070

Solid line = Upper bound favorites of 18.5 points or less; N = 29,723

Shaded area is a reference band; the result of a bootstrap hypothesis test of equality.

Test of equal densities: p-value = 0

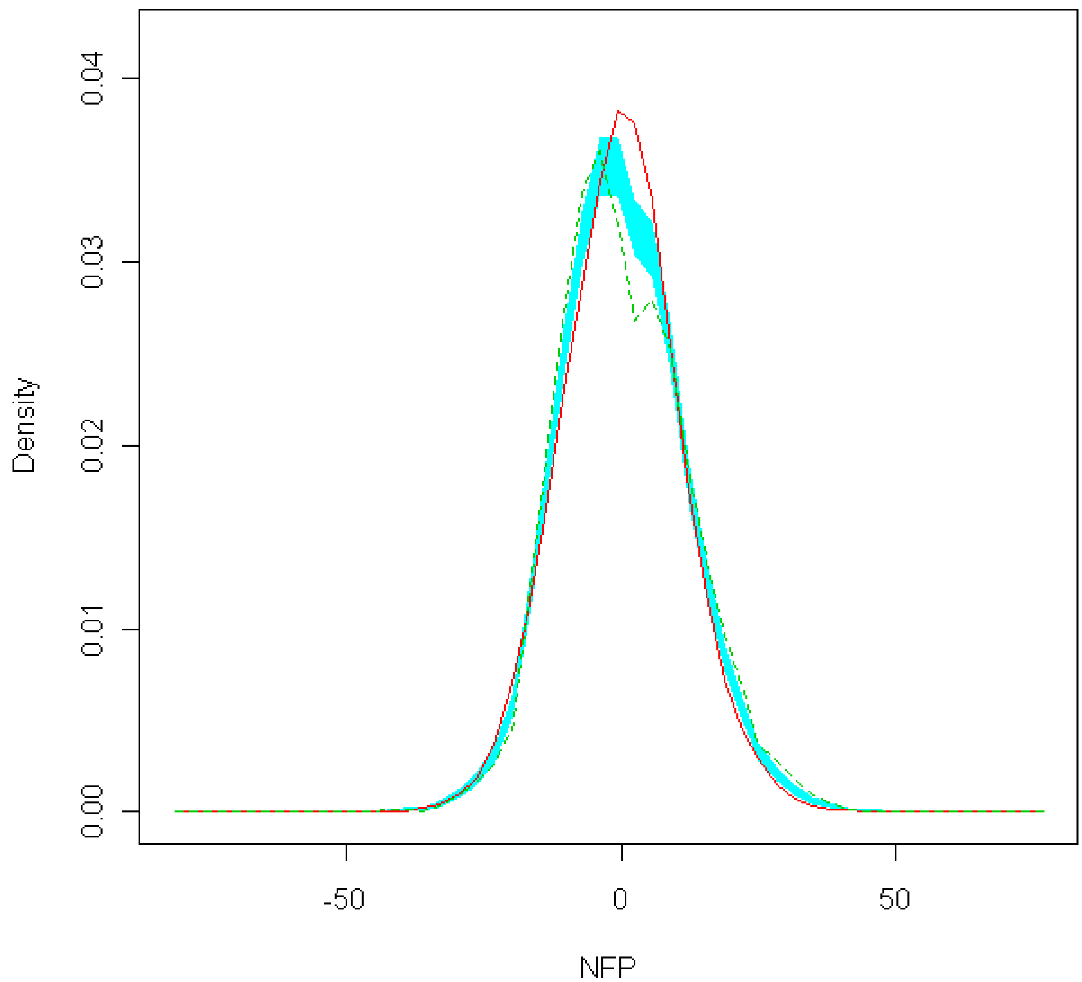

Figure 3: Regular Season Test of two distributions: 11.5 Point Favorites

Dashed line = Lower Bound favorites of 11.5 points or more; N = 8,440
Solid line = Upper bound favorites of 11 points or less; N = 23,353
Shaded area is a reference band; the result of a bootstrap hypothesis test of equality.
Test of equal densities: p-value = 0

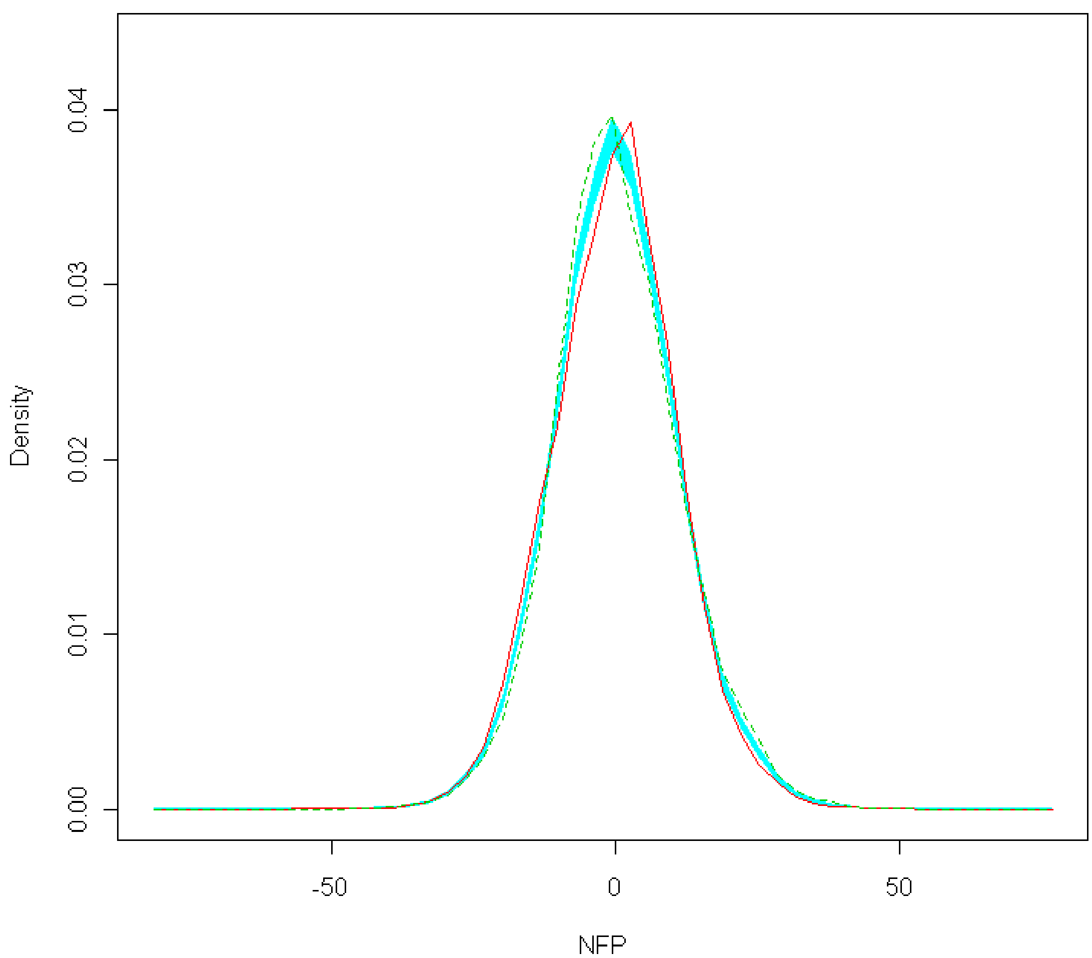

Figure 4: Regular Season Test of two distributions: 3.5 Point Favorites

Dashed line = Lower Bound favorites of 3.5 points or more; N = 24,730

Solid line = Upper bound favorites of 4 points or less; N = 7,063

Shaded area is a reference band; the result of a bootstrap hypothesis test of equality.

Test of equal densities: p-value = 0

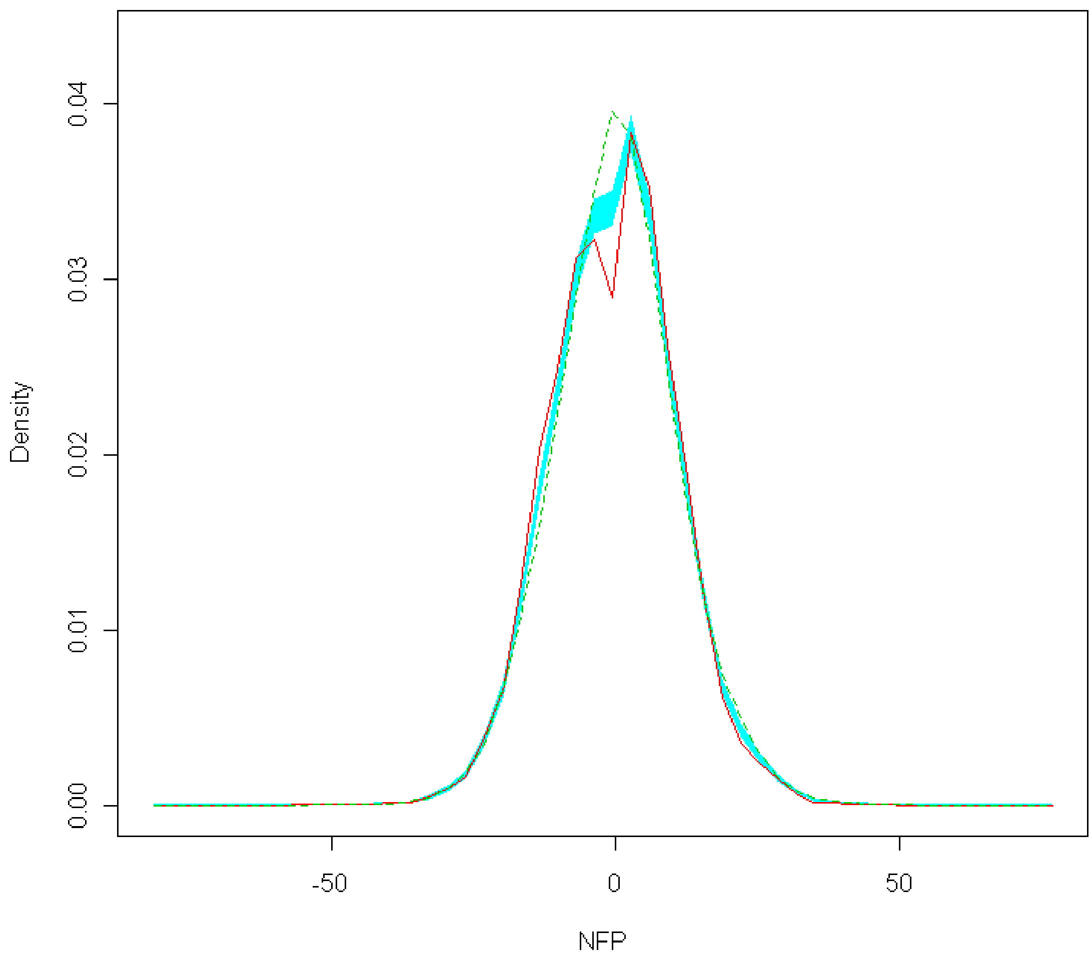

Figure 5:Probability Distribution: Favored Outcomes Relative to Point Spread Tournament Play

“x” axis measures the favored team’s outcome relative to the point spread (NFP).

N = 3,371

Shaded area represents a normal distribution reference band

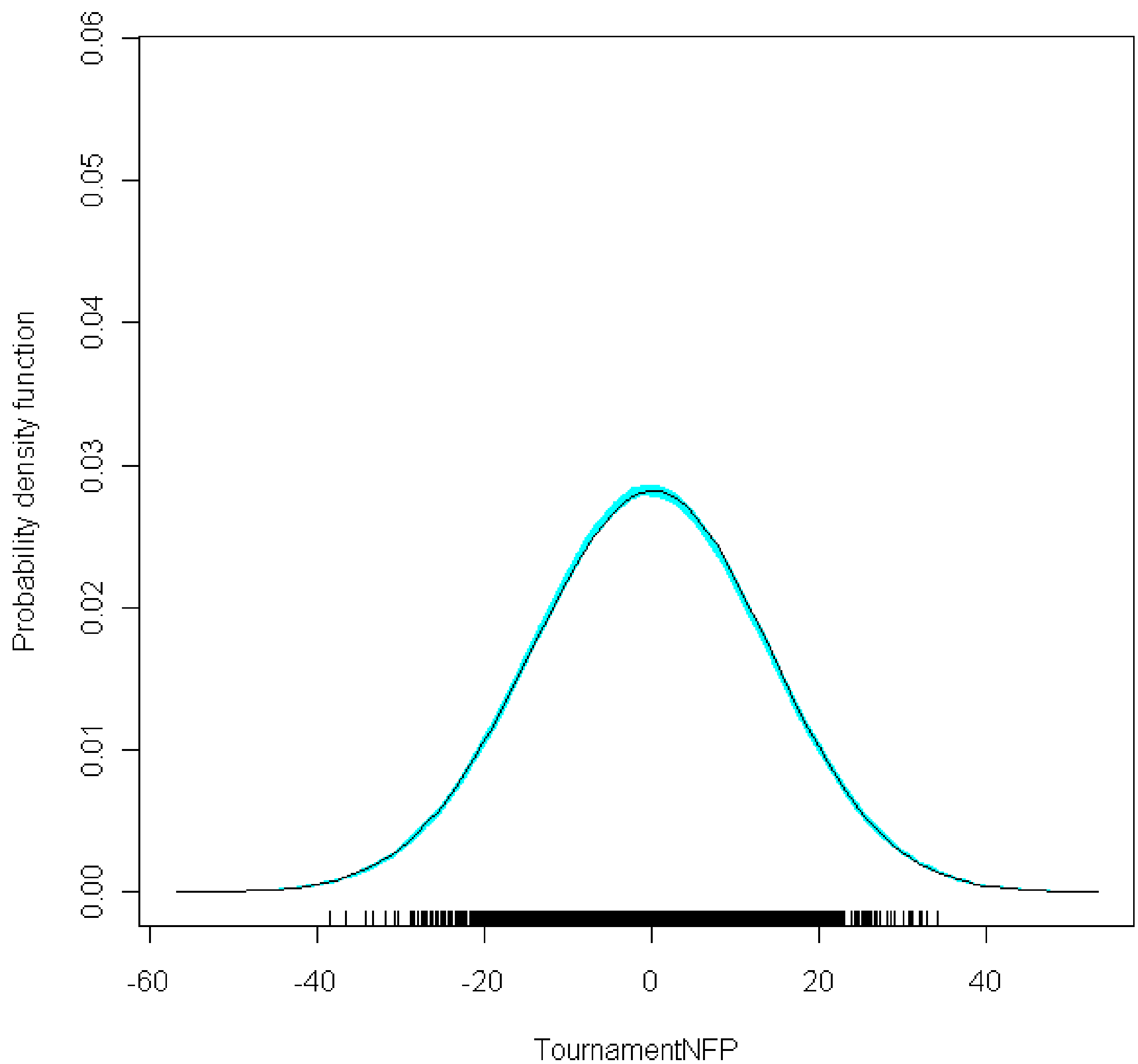

Figure 6: Tournament Play Test of two distributions: 19 Point Favorites

Dashed line = Lower Bound favorites of 19 points or more; N = 99

Solid line = Upper bound favorites of 18.5 points or less; N = 3,272

Shaded area is a reference band; the result of a bootstrap hypothesis test of equality.

Test of equal densities: p-value = 0.28

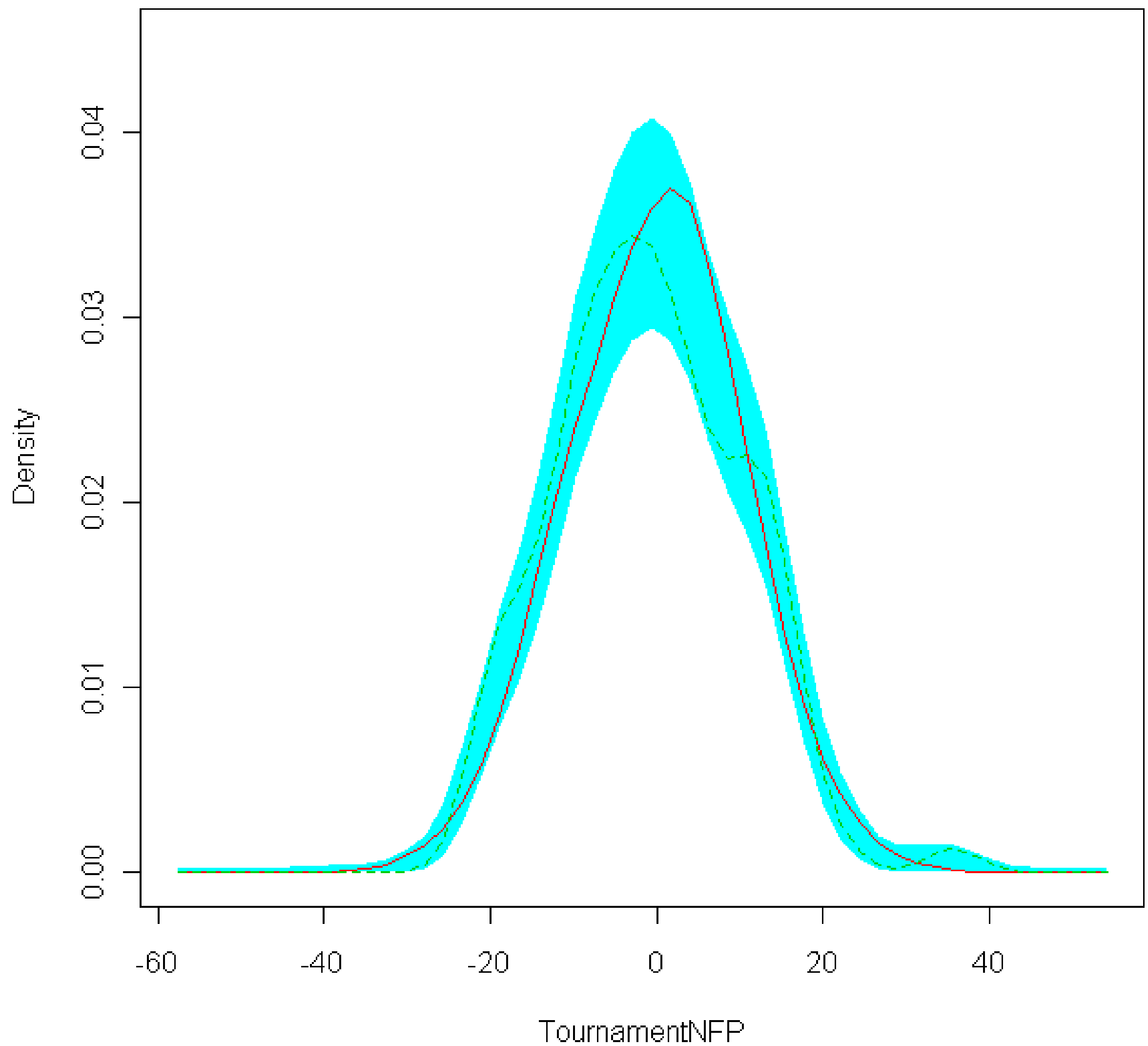

Figure 7: Tournament Play Test of two distributions: 11.5 Point Favorites

Dashed line = Lower Bound favorites of 11.5 points or more; N = 520
Solid line = Upper bound favorites of 11 points or less; N = 2,851
Shaded area is a reference band; the result of a bootstrap hypothesis test of equality.
Test of equal densities: p-value = 0.35

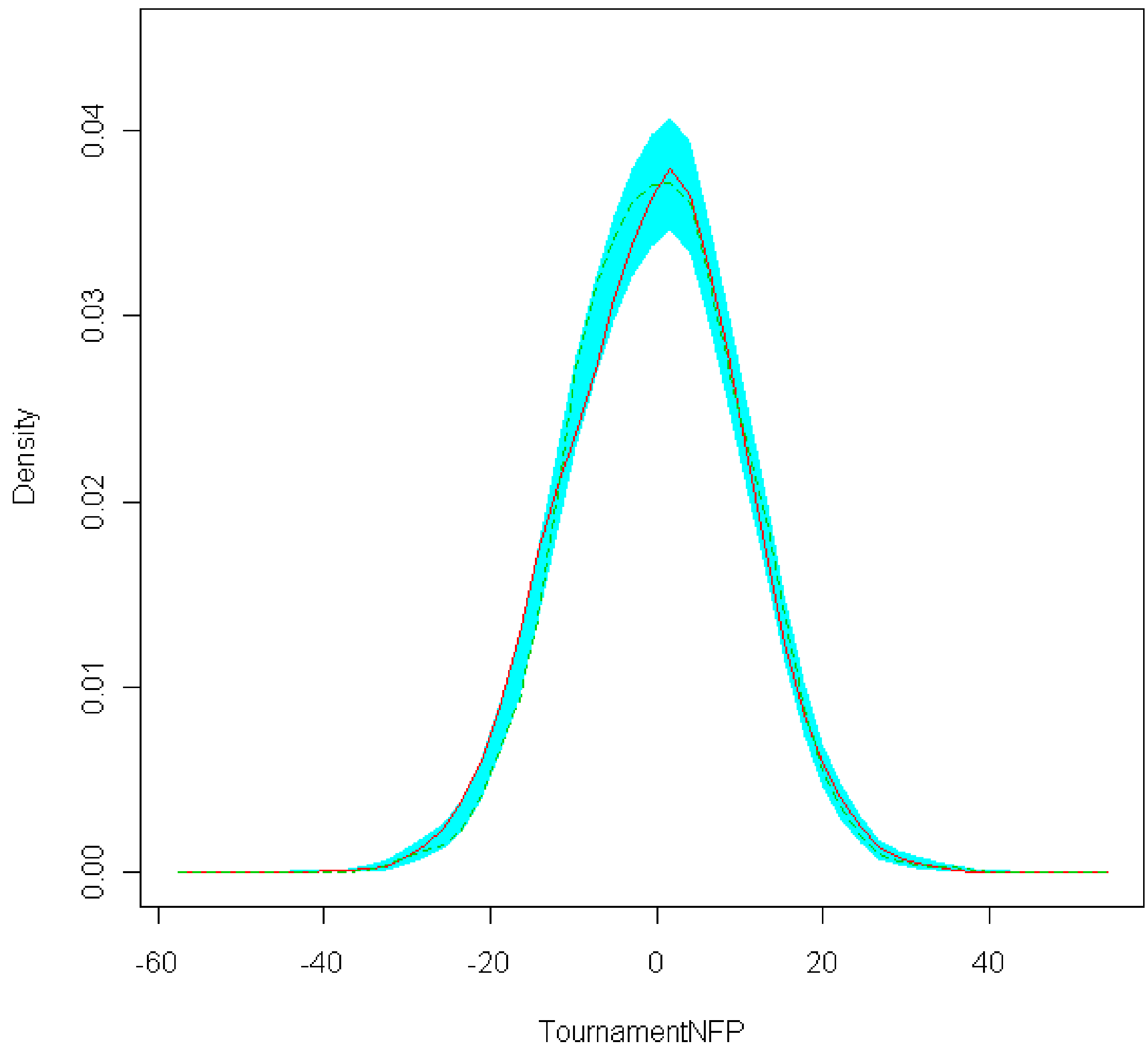

Table 1: Regression of Point Differential on Point Spread			
Coefficient	Wolfers	Regular Season	Tournament
Point Spread	1.007 (167.83)	1.048*** (104.23)	1.045*** (29.92)
Constant	-0.012 (0.20)	-0.406*** (3.94)	0.503* (1.70)
Adjusted R-squared	0.39	0.26	0.21
Breusch-Pagan Chi-squared	N/A	8.17***	0.83

t-statistics in parentheses

**Significant at 10-percent level*

***Significant at 5-percent level*

****Significant at 1-percent level*

Sources

Adrian Bowman and Adelchi Azzalini, *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. (London, Oxford University Press, 1997)

Berhardt, Dan and Steven Heston. (2010) "Point Shaving in College Basketball: A Cautionary Tale for Forensic Economics," *Economic Inquiry,* 48(1), January, pp. 14-25.

Borghesi, Richard (2008), "Widespread Corruption in Sports Gambling: Fact or Fiction?" *Southern Economic Journal 74(4)* at p1063-1069

Brown, Robert. (1994) "Measuring Cartel Rents in the College Basketball Recruitment Market," *Applied Economics*, 26(1), January, pp. 27-34.

Diemer, George, (2009A) "Internet Sports Books Operating in an Inefficient Marketplace," *The Journal of Gambling Business and Economics 3(1)* at p1-22

Diemer, George. (2009B) "Point Shaving in the NFL Gambling Market: A Bootstrap Investigation," *Journal of Prediction Markets,"* 3(3), December, pp. 13-31.

Johnson, Neal. (2009) "NCAA Point Shaving As An Artifact of the Regression Effect and the Lack of Tie Games," *Journal of Sports Economics,* 10(1), February, pp. 59-67.

Leeds, Michael A. and Peter von Allmen. (2010) *The Economics of Sports,* Boston: Pearson.

Michael McCarthy, (2007) "Point Shaving Remains a Concern in College Athletics," *USA Today*, May 9, on-line at http://www.usatoday.com/sports/college/2007-05-08-point-shaving-cover_N.htm.

Seymour, Harold. (1960) *Baseball: The Early Years,* Oxford: Oxford University Press.

Vaughan Williams, L., Ed. (2005) *Information Efficiency in Financial and Betting Markets.* Cambridge: Cambridge University Press.

Wolfers, Justin. (2006) "Point Shaving: Corruption in NCAA Basketball" *American Economic Review, 96(2),* t p279-283

Wolfers, Justin and Eric Zitzewitz. (2004) "Prediction Markets" *Journal of Economic Perspectives, 18(2),* p107-26