The NCAA Basketball Betting Market: Tests of the Balanced Book and Levitt Hypotheses

Rodney J. Paul, St. Bonaventure University

Andrew P. Weinbach, Coastal Carolina University

Kristin K. Paul, St. Bonaventure University

Abstract

Sportsbook behavior is tested for NCAA basketball using actual sportsbook betting percentages from on-line sportsbooks. The balanced book hypothesis of the traditional sportsbook models does not appear to hold, as favorites attract more than 50% of the bets. Although there is some slight evidence toward shading the line in these directions, there is also not overwhelming evidence of the Levitt (2004) hypothesis, as sportsbooks do not appear to be actively pricing to maximize profits. In general, the results seem more consistent with the sportsbook pricing as a forecast, content with earning their commission on losing bets as simple strategies win about 50% of the time.

Introduction

A study by Levitt (2004) in *The Economic Journal* challenged the traditional view of sportsbook behavior. In the Levitt hypothesis, sportsbooks set prices to maximize profits, not to balance the sports betting action. This model differs substantially from the traditional models of sportsbook behavior, such as Pankoff (1968), Zuber, et. al. (1985), and Sauer, et. al. (1988), where sportsbooks set prices to balance the book. They achieve this by setting a price which attracts equal dollars on each side of the betting proposition. Under this model, using sports betting data to test the efficient markets hypothesis is straightforward. Under the assumptions of the traditional models, the efficient markets hypothesis could be tested with relative ease as the price represents information from all betting participants. Findings that the efficient markets hypothesis could not be rejected, even in a market where investor (bettor) sentiment is likely to run high, served as a measure of support for this theory (i.e. Sauer, et. al. 1988).

If sportsbooks are not pricing to balance the book, however, comparisons between sports wagering markets and other financial markets (such as stocks or bonds), particularly in the testing of the efficient markets hypothesis, become suspect. If prices are being set by sportsbooks to maximize profits or are set as a forecast of game outcomes, independent of the flow of betting dollars, prices in these markets are no longer formed by the actions of investors (bettors), but by the sportsbook itself.

One common criticism of the empirical findings of Levitt (2004) is the use of a betting tournament to substantiate the theory, rather than use of actual sportsbook data. The tournament in question used a limited number of participants with a fixed entry fee of $250. The results from this tournament could yield vastly different results from an actual sportsbook, which has a large number of participants who place wagers of varying sizes on games they bet.

In a recent article in the *Journal of Prediction Markets*, Paul and Weinbach (2007) used actual sportsbook data to test the hypothesis of Levitt (2004) concerning sportsbook behavior. Actual percentages of dollars wagered on the favorite and the underdog were obtained for every game of the 2006 NFL season. The results for the pointspread market were consistent with the results of Levitt (2004), as betting did not appear to be balanced, with favorites, in particular road favorites, receiving a greater percentage of betting volume. In addition, the percentage bet on the favorite became greater as the pointspread on the favorite increased. Simple strategies of betting against the public, when the sportsbook was substantially unbalanced (i.e. 70%+ on the favorite) were found to earn positive returns. Similar findings concerning an unbalanced book and bettor preferences for favorites and overs were found in the NBA (Paul and Weinbach, 2008), although evidence of the hypothesis of Levitt concerning pricing to maximize profits was not found.

This paper explores the wagering market for college basketball, using the same data source used by Paul and Weinbach (2007, 2008). Tests of the traditional model of sportsbook behavior compared to the findings of Levitt (2004) are performed. The traditional model of sportsbook behavior assumes that sportsbooks set prices to balance the book. Therefore, under the traditional model, prices (pointspreads) are assumed to move based upon actions of bettors. If bettors wager more on the favorite, for example, the pointspread is expected to increase. The Levitt Hypothesis assumes that the sportsbook is not balanced, but the sportsbook uses their knowledge of the behavioral biases of bettors to set prices to maximize profits.

Regression results illustrating the relationship between the pointspread and the percentage bet on the favorite are shown. Betting simulations are also presented to test if the sportsbook purposefully allows a betting imbalance to maximize profits. In addition, our hypothesis that sportsbooks price as a forecast of the outcome of the game, independent of the actions of bettors, is explored.

Regression and Betting Simulation Results

Data for this paper were gathered from Sports Insights, which sells data to subscribers including the percentage of bets made on each proposition within each game. These data includes the percentage bet on favorites and underdogs in the pointspread market. Sports Insights presents combined data from four sportsbooks to show the percentage of bets on the favorite and underdog for its subscribers. The four on-line sportsbooks are BetUS.com, CaribSports.com, SportBet.com, and Sportsbook.com. Sports Insights reports percentages based on the *number of bets* placed on each side of the proposition. The number of bets is not a perfect measure, as bets do vary in magnitude; however, the number of bets presented in the Sports Insights data appears proportionally similar to the dollars wagered volume in the Sportsbook.com data.

A simple regression model is tested, which illustrates the actions of the sportsbook. The model to be estimated is as follows for the sides (pointspread) market:

$$(\text{\% Bet on the Favorite})_i = \alpha_0 + \beta_1(\text{Pointspread})_i + \beta_2(\text{Dummy for Road Favorite})_i + \varepsilon_i \quad (1)$$

The dependent variable is the percentage of dollars bet on the favorite. The independent variables include an intercept, the pointspread on the game (presented as a positive number – greater favorites have larger pointspreads), and a dummy for teams which are road favorites. If bettors prefer favorites, with stronger favorites being bet more heavily than weaker favorites, the coefficient β_1 should be positive and significant. If bettors overbet road favorites, the coefficient on the dummy variable, β_2, should also be positive and significant.

Table 1 presents the results for the pointspread market for NCAA Basketball. Coefficients on the independent variables are shown, with t-stats in parentheses. Heteroskedasticity was found in the initial regression results, therefore White's heteroskedasticity-consistent standard errors and covariances were used and are presented in the table below.

Table 1: NCAA Sides Regression 2004-05 to 2006-07
Dependent Variable: Percentage of Bets on the Favorite
Number of Observations: 12,644

Independent Variables	Coefficient (T-Statistic)
Constant	56.0024*** (193.7622)
Pointspread	0.4288*** (15.7621)
Road Favorite Dummy	10.2101*** (28.1731)

From the results in Table 1, it appears the results for NCAA basketball are similar to the results in the NFL (Paul and Weinbach, 2007) in relation to the percentage of bets placed on favorites. As the pointspread on the favorite increases, the percentage of bets on the favorite also increases, by 0.4288 percentage points for each additional point on the pointspread. From the regression model, a 10-point home favorite is expected to attract over 60% of the betting action. A 20-point favorite is expected to attract over 64% of the betting action.

Similar to what is described in Levitt (2004), road favorites are also found to be significantly overbet, as the dummy variable for a road favorite is positive and significant. An additional 10%+ of the bets accumulate on the favorite when the favorite is playing on the road. NCAA basketball bettors seem to prefer to wager on the best teams, given the significance and the positive signs found on the pointspread variable and the road favorite dummy.

Given the balanced book hypothesis can be rejected, as bets on favorites are not found to be 50% across the sample, the next step is to determine if sportsbooks set prices (pointspreads) to maximize profits by exploiting known bettor biases for favorites, and in particular, road favorites. An alternative explanation to the balanced book hypothesis and the hypothesis of Levitt, that the sportsbook is pricing as a forecast, is also explored. First, however, basic market efficiency and returns to simple betting strategies are shown.

Betting Simulations of Wagering on Underdogs in NCAA Basketball

Market efficiency has previously been studied for the NCAA Basketball betting market. Paul and Weinbach (2005a) found that the overall market for college basketball appeared efficient, but wagering on a simple strategy of betting big underdogs (defined as double-digit underdogs), and especially home underdogs, was found to reject market efficiency and generate profitable returns. These results were found to be similar to other sports such as college football (Paul and Weinbach, 2003), and the NBA (Paul and Weinbach, 2005b). Wolfers (2006) also showed a similar bias of heavy underdogs winning more than implied by efficiency (defined in his sample as 12 or more point underdogs) and attracted attention with his allegations of pointshaving as the source of this bias.

Tables 2-4 present the results of simple betting simulations of wagering on underdogs in college basketball. Given the results shown in Table 1, the higher the pointspread on the game, the greater the percentage of the bets on the favorite. Therefore, we show the results of wagering on underdogs (the less popular side of the proposition), when they meet certain thresholds, for various categories (ten points or greater, eight points or greater, etc.) and for all games. Results are shown for all favorites, all home favorites, and all road favorites. For each category, favorite wins, underdog wins, the underdog win percentage, and the log likelihood ratio test of a fair bet are shown.

In tables 2-4, none of the win percentages based on these simple strategies could reject the null of no profitability (and infrequently have win percentages greater than 52.38%, the win percentage required to break even), therefore only tests for the null of a fair bet (win percentage equals 50%) are shown. Significant results are noted for the log likelihood ratio tests with * representing significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 2: Betting Simulations for All Underdogs – Strategy of Bet the Underdog

All Favorites Greater Than:	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log Likelihood Ratio Test: Fair Bet
20	210	229	52.1640%	0.8226
18	317	355	52.8274%	2.1500
16	490	545	52.6570%	2.9241*
14	752	825	52.3145%	3.3804*
12	1173	1254	51.6687%	2.7038
10	1660	1743	51.2195%	2.0246
8	2271	2420	51.5881%	4.7335**
6	3131	3226	50.7472%	1.4197
4	4077	4222	50.8736%	2.5336
2	5221	5312	50.4320%	0.7862
All	6122	6250	50.5173%	1.3243

The log likelihood test statistics have a chi-square distribution with one degree of freedom. Critical Values are 2.706 (for an α=0.10), 3.841 (for an α=0.05), and 6.635 (for an α=0.01). * is significance at 10%, and ** is significance at 5%.

Table 3: Betting Simulations for All Road Underdogs (Home Favorites) – Strategy of Bet the Underdog

All Home Favorites Greater Than:	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log Likelihood Ratio Test: Fair Bet
20	202	222	52.3585%	0.9437
18	302	343	53.1783%	2.6080
16	465	525	53.1783%	3.6386*
14	708	779	52.3874%	3.3913*
12	1085	1161	51.6919%	2.5722
10	1514	1597	51.3340%	2.2147
8	2015	2170	51.8519%	5.7421**
6	2701	2802	50.9177%	1.8538
4	3365	3544	51.2954%	4.6381**
2	4123	4270	50.8757%	2.5748
All	4635	4794	50.8431%	2.6813

The log likelihood test statistics have a chi-square distribution with one degree of freedom. Critical Values are 2.706 (for an α=0.10), 3.841 (for an α=0.05), and 6.635 (for an α=0.01). * is significance at 10%, and ** is significance at 5%.

Table 4: Betting Simulations for All Home Underdogs (Road Favorites) – Strategy of Bet the Underdog

All Road Favorites Greater Than:	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log Likelihood Ratio Test: Fair Bet
20	8	7	46.6667%	0.06672
18	15	12	44.4444%	0.3340
16	25	20	44.4444%	0.5567
14	44	46	51.1111%	0.0444
12	88	93	51.3812%	0.1381
10	146	146	50.0000%	0.0000
8	256	250	49.4071%	0.0711
6	430	424	49.6487%	0.04215
4	712	678	48.7770%	0.8317
2	1098	1042	48.6916%	1.4656
All	1487	1456	49.4733%	0.3254

The log likelihood test statistics have a chi-square distribution with one degree of freedom. Critical Values are 2.706 (for an α=0.10), 3.841 (for an α=0.05), and 6.635 (for an α=0.01). * is significance at 10%, and ** is significance at 5%.

For the sample of all favorites (table 2), games with larger favorites tend to have the underdog cover the pointspread more than fifty percent of the time. These percentages are not high enough (52-53%) to reject the null of no profitability, but betting the underdog for all favorites greater than 8 was found to reject the null of a fair bet at the 5% level. Betting all underdogs greater than 14 and 16 were also found to reject the null hypothesis of a fair bet at the 10% level.

Table 3 provides the results for road underdogs (home favorites). For this sample, rejections of the null hypothesis of a fair bet were found at the 5% level for the subset of all underdogs greater than 4 and 8. Rejections of the null of a fair bet were also found at the 10% level for underdogs greater than 14 and 16. Table 4 shows the results for home underdogs (road favorites). The null hypothesis of a fair bet could not be rejected for this sample or any of its sub-groupings.

Overall, underdogs win slightly more than 50% of the games. These win percentages, however, do not generally generate profits for underdog bettors. Although favorites, particularly road favorites and big favorites, attract a greater share of the betting action, the closing pointspreads do not appear to be greatly biased, as simple strategies of wagering on the underdog in these situations does not generate statistically significant profits.

Betting Simulations of Wagering Against Public Sentiment

Betting against public sentiment may also be a possible winning strategy in the betting market for college basketball. If large betting imbalances illustrate preferences of bettors for favorites, perhaps sportsbooks respond by shading the pointspread in the direction of this sentiment. This appears to be the case in the NFL (Paul and Weinbach, 2007) and it is useful to know if it also exists in NCAA basketball.

If sportsbooks are setting pointspreads (prices) to maximize profits, as suggested by Levitt (2004), a simple contrarian strategy of placing a wager on the side of the proposition in which the sportsbook is exposed, specifically, wagering on the publically unpopular underdogs, should generate positive returns. If this is not the case, the sportsbook would not be pricing to maximize profits by exploiting known betting biases (such as road favorites in the NFL).

Table 5 shows the results of betting against public sentiment. Results are shown based on a simple strategy of betting against the public in games where the sportsbook is heavily weighted (greater than 80%, greater than 70%, etc.) on the favorite. Win percentages of a simple strategy of bet the underdog is shown along with the log likelihood ratio test for the null of a fair bet.

Table 5: Betting Simulations of Betting the Opposite of Public Sentiment

Percentage Bet on Favorite is Greater Than:	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log-Likelihood Ratio Test: Fair Bet
80%	62	65	51.1811%	0.0709
70%	223	215	49.0868%	0.1461
60%	603	631	51.1345%	0.6354
50%	1350	1384	50.6218%	0.4228
All	6123	6251	50.5172%	1.3241

The log likelihood test statistics have a chi-square distribution with one degree of freedom. Critical Values are 2.706 (for an α=0.10), 3.841 (for an α=0.05), and 6.635 (for an α=0.01). * is significance at 10%, and ** is significance at 5%.

Betting against public sentiment does not appear to be a profitable venture in NCAA basketball. Win percentages of these strategies tend to hover around 50% for any chosen threshold. No matter how large the imbalance of bets, wagering against (or with) the public money tends to leave the bettor winning about half of his bets and losing the commission on these bets over time.

These results, coupled with the large betting imbalances on NCAA Basketball games, support the notion that sportsbooks are not attempting to balance the betting dollars on favorites and underdogs, as commonly assumed by the traditional models of sportsbook behavior. There is some evidence the largest favorites are overpriced, as simple betting strategies of betting big underdogs win more often than implied by efficiency, although these win percentages are not significant compared to the null hypothesis of no profitability. In addition, betting against the largest betting imbalances toward the favorite is not found to win often enough to reject the null of a fair bet. This is in contrast with the findings of Levitt (2004) and Paul and Weinbach (2007) for the NFL.

It is possible that the size of the market plays a significant role in whether sportsbooks will attempt to price to maximize profits by exploiting bettor biases. The normal NCAA basketball game is a much less popular betting proposition than the average NFL game, and there are many more NCAA basketball games per season. Therefore, sportsbooks may not be as willing to set prices to attempt to take advantage of common bettor biases in college basketball. Instead, they may attempt to set pointspreads as optimal and unbiased predictors of outcomes of games.

Setting of the pointspread as a forecast, independent of the betting dollar percentages on favorites and underdogs, is consistent with the results, where each side of the proposition wins approximately 50% of the time. Under the assumption of betting as a repeated game, over the course of a season or many seasons, the sportsbook may be content to set the pointspread as a forecast of the outcome of the game. Instead of possibly inviting informed bettors into the fold by inflating prices (pointspreads) on big favorites, the sportsbook may be content to price with the expectations of favorites and underdogs each winning half of the time.

Conclusions

The betting market for NCAA basketball was tested in relation to sportsbook pricing behavior using actual betting percentages from real sportsbooks. The results of these tests were compared to previous results found on betting percentages in the NFL (Paul and Weinbach, 2007) and the NBA (Paul and Weinbach, 2008). Using the betting percentages on each game, support was attempted to be found for the traditional models of sportsbook behavior, where the book is balanced, the Levitt hypothesis, where sportsbooks price to maximize profits, or a hybrid model where sportsbooks price as a forecast, allowing an unbalanced book, but not exploiting known bettor biases to maximize profits.

In general, the traditional model of sportsbook behavior does not appear to be supported as the betting dollars in college basketball are not balanced. Favorites and overs tend to attract a higher percentage of the betting action. These results do not necessarily imply that sportsbooks are pricing to exploit known biases and maximize profits, as Levitt (2004) suggests.

To test if sportsbooks price to maximize profits by exploiting known bettor biases, a couple of simple tests were performed on the data. First, simple betting strategies of betting the underdog and the under were performed. Underdogs won slightly more often than favorites, but the results were not found to be statistically significant in the sample of all games. For all underdogs of 8 or more, 14 or more, and 16 or more, however, statistical significance was found.

When considering betting percentages and calculating the results when the betting public is heavy on the favorite or over (meaning the sportsbook is an active participant on the side of the underdog or under), little in the way of statistical significance was found. The only case where a fair bet could be rejected was in situations where the public had 70% or more on the favorite in an AFL game, where the underdog won more than implied by efficiency. The rest of the results of these tests could not reject the null of a fair bet. Betting on road underdogs of 4 or more, 8 or more, 14 or more, and 16 or more were also found to win more than implied by efficiency. None of the groupings for home underdogs were found to have significant results. In addition, using the betting percentages and wagering against public sentiment (or with public sentiment) was not found to generate significant returns.

Overall it appears there may be some slight shading of the pointspread toward the favorite, but the returns for the overall sample are not enough to generate profitable returns. Similar to other studies, however, some groupings of large underdogs are shown to have statistically significant returns. Given that the betting action is not found to be balanced, but profitability is not found to a great extent by taking the side of the sportsbook (underdogs and unders), it appears that the sportsbook does not follow the traditional model of sportsbook behavior nor the Levitt hypothesis. It appears that sportsbooks price generally as a forecast, with a slight shade (particularly in obvious cases – big favorites or high totals) toward the more popular side of the proposition.

These findings for NCAA basketball are more similar to the NBA (Paul and Weinbach, 2008) than the NFL (Paul and Weinbach, 2007). We believe the NCAA basketball results are closer to the NBA results due to both sports being less popular with bettors, per game, than NFL football. NFL football consistently attracts the most bettors and betting dollars by a wide margin compared to other sports. Given the large market for NFL football betting, it may be in the interest of the sportsbook to shade the pointspreads to attempt to earn greater profits. This strategy likely results in higher transactions costs for the sportsbook as it must more closely monitor the wagering activity on each game, as informed bettors (if they exist) may exploit inflated lines. To prevent this, sportsbooks are likely to practice what is called "booking to face" when bettors are treated heterogeneously with suspected informed bettors facing lower betting limits and/or refused wagers outright. These strategies are costly for the sportsbook, but could be worth it due to the level of betting action seen in the NFL, which is normally not present in other sports, such as college and professional basketball.

References

Levitt, S. (2004). Why are gambling markets organized so differently from financial markets? *The Economic Journal*, 114, 223-46.

Pankoff, Lyn (1968). Market Efficiency and Football Betting. *Journal of Business*, 41, 203-214.

Paul, Rodney J., Weinbach, Andrew P., and Weinbach, Chris. (2003). Fair Bets and Profitability in College Football Gambling. *Journal of Economics and Finance.* 27:2, 236-242.

Paul, Rodney J. and Weinbach, Andrew P. (2005a). Market Efficiency and NCAA College Basketball Gambling. *Journal of Economics and Finance* 29(3): 403-408.

Paul, Rodney J. and Weinbach, Andrew P. (2005b). Bettor Misperceptions in the NBA: The Overbetting of Large Favorites and the "Hot Hand," *Journal of Sports Economics* 6(4), 390-400.

Paul, R. and Weinbach, A. (2007). Does Sportsbook.com Set Pointspreads to Maximize Profits? Tests of the Levitt Model of Sportsbook Behavior. *Journal of Prediction Markets,* 1(3), 209-218.

Paul, R. and Weinbach, A. (2008). Price Setting in the NBA Gambling Market: Tests of the Levitt Model of Sportsbook Behavior. *International Journal of Sports Finance*, 3, 3, 2-18.

Sauer, R., Brajer, V., Ferris, S. and Marr, M. (1988). Hold your bets: Another look at the efficiency of the gambling market for National Football League games. *Journal of Political Economy*, 96, 206-13.

Wolfers, J. (2006). Pointshaving: Corruption in NCAA Basketball. *American Economic Review*, 96(2), 279-283.

Zuber, R., Gandar, J. and Bowers, B. (1985). Beating the spread: Testing the efficiency of the gambling market for National Football League games. *Journal of Political Economy*, 93, 800-6.