

Using Actual Betting Percentages to Analyze Sportsbook Behavior: The Canadian and Arena Football Leagues

Rodney J. Paul*, Andrew P. Weinbach**, and Kristin K. Paul***

ABSTRACT

Sportsbook behavior is tested for the Canadian and Arena Football Leagues using real sportsbook betting percentages from on-line sportsbooks. The balanced book hypothesis of the traditional sportsbook models does not appear to hold for these leagues, as favorites and overs attract more than 50 percent of the betting dollars. Although there is some slight evidence toward shading the line in these directions, there is also no overwhelming evidence supporting 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 percent of the time.

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 that 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 (e.g. 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

*229 Murphy Building, School of Business, Saint Bonaventure University, Saint Bonaventure, NY 14778, (716) 375-2145, rpaul@sbu.edu

**Coastal Carolina University

***St. Bonaventure University

FALL 2010

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 the games on which they bet.

In a recent article in the *Journal of Prediction Markets*, Paul and Weinbach (2007) used actual sportsbook data to test Levitt's (2004) hypothesis. 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 the 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 percent-plus on the favorite) were found to earn positive returns.

Paul and Weinbach (2007) also showed similar results for the totals (over/under) betting market for the NFL as sportsbook.com was found to be unbalanced, with bettors heavily weighted toward the over, which was consistent with the results seen for long samples of total betting in the NFL (Paul and Weinbach, 2002). Similar findings concerning an unbalanced book and bettor preferences for favorites and overs were found in the NBA (Paul and Weinbach, 2008).

This paper explores the wagering market for smaller betting markets, the Canadian Football League and the Arena Football League, employing 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. 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, the possibility that the sportsbooks' price is a forecast of the outcome of the game, independent of the actions of bettors, is explored.

II. Regression and Betting Simulation Results

Data from Sports Insights, which provides actual betting information from on-line sportsbooks for the Canadian Football League (CFL) and Arena Football League (AFL), were purchased from their website, www.sportsinsights.com. Data were gathered for both the sides (betting on a team against the pointspread) and totals (betting on the total amount of points scored by both teams) markets. Following the method used by Paul and Weinbach for the NFL (2007) and the NBA (2008), we set up a regression to illustrate how the percentage bet on the favorite and the over vary with the magnitude of the pointspread and total and how the existence of a road favorite affects the betting percentages.

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

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

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 since heavier favorites have larger pointspreads), and a dummy for teams which are road favorites. Road favorites have been shown to be commonly overbet in studies such as Golec and Tomarkin (1991) and Gray and Gray (1997). These studies were cited and used to study the betting tournament data in Levitt (2004). The road favorite dummy variable was found to be positive and significant for the NFL (Paul and Weinbach, 2007) and NBA (Paul and Weinbach, 2008).

A couple of simple propositions can be tested from this regression model. First, if bettors overbet favorites and stronger favorites are 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.

The totals market is tested in the same manner as the sides market. The simple regression model for the totals market is:

$$(\% \text{ Bet on the Over})_i = \alpha_0 + \beta_1(\text{Total})_i + \varepsilon_i. \quad (2)$$

If more wagers are accepted on the over as total increases, then β_1 should be positive and significant. Market efficiency is tested through the joint null hypothesis that the intercept is zero and the coefficient on the independent variable (either the pointspread or total) is equal to one. The results for the sides (pointspread) regression are presented in tables 1 (2005-2008 CFL) and 2 (2005-2007 AFL). The results for the totals (over/under) regression are presented in tables 3 (2005-2008 CFL) and 4 (2005-2007 AFL).

Table 1: Pointspread Betting Percentages – Sports Insights 2005-2007 CFL

Dep. Var: Percent Bet on Favorite	Coefficient	Standard Error	T-Statistic
Constant	59.9959	2.3811	25.1970***
Pointspread	-0.6416	0.3362	-1.9081*
Road Favorite Dummy	1.1337	2.4249	0.4674

*-notation denotes statistical significance at the following levels - *-10%, **-5%, and ***-1%.

Table 2: Pointspread Betting Percentages – Sports Insights 2005-2008 AFL

Dep. Var: Percent Bet on Favorite	Coefficient	Standard Error	T-Statistic
Constant	52.2719	1.2928	40.4329***
Pointspread	0.1975	0.1519	1.2409
Road Favorite Dummy	9.0638	1.3446	6.7409***

*-notation denotes statistical significance at the following levels - *-10%, **-5%, and ***-1%.

Table 3: Totals Betting Percentages – Sports Insights 2005-2007 CFL

Dep. Var: Percent Bet on Over	Coefficient	Standard Error	T-Statistic
Constant	13.4582	15.7678	0.8535
Total	0.7723	0.3059	2.5251**

*-notation denotes statistical significance at the following levels - *-10%, **-5%, and ***-1%.

Table 4: Totals Betting Percentages – Sports Insights 2005-2008 AFL

Dep. Var: Percent Bet on Over	Coefficient	Standard Error	T-Statistic
Constant	24.7974	9.4179	2.6330***
Total	0.2966	0.0893	3.3222***

*-notation denotes statistical significance at the following levels - *-10%, **-5%, and ***-1%.

F-tests were performed for each regression result. The F-test of whether the intercept equaled 50 and the pointspread (total) variable was equal to zero was tested for both pointspreads and totals, while additional F-tests of the intercept equaling 50, the pointspread variable equaling zero, and the road favorite dummy equaling zero were performed for the pointspread regressions. In each case, the null hypothesis could be rejected at the 1 percent level. The F-test results for whether the intercept equaled 50 and the coefficient on the pointspread (total) was zero were found to be (F-values) 47.007 (AFL-Pointspread), 7.5063 (AFL-Totals), 17.571 (CFL-Pointspread), and 6.614 (CFL-Totals).

In tables 1 and 2, the intercept of both the CFL and AFL regression is greater than 50, implying that on the average more than half of the bets accrue to the favorite. The CFL, however, does not look like the NFL, as bigger favorites do not receive an increasing share of the betting dollars, but actually receive slightly less than smaller favorites (significant at the 10 percent level). The road favorite dummy for the CFL is positive, but not statistically significant.

In the AFL, bets on the favorite increase with each additional point of the pointspread, as in the NFL and NBA, but this variable is not found to be statistically significant for the Arena League. The road favorite dummy variable is found to have a positive and significant effect on the amount bet on the favorite. This is similar to the result found in the NFL and NBA, as bettors appear to favor road favorites, which likely reveals a strong preference for the best teams, as it takes a good team to be a road favorite.

Overall, in the sides regressions for the CFL and AFL, some preference for the favorite is found, but the preference for the biggest favorites is not as great as in the NFL and NBA. Road favorites still appear to be popular in the CFL and AFL, although statistical significance is only found for the AFL.

Tables 3 and 4 present the results for the totals (over/under) market. Strong preference for the over was found in the NFL and the NBA (Paul and Weinbach, 2007, 2008) and CFL and AFL bettors appear to have the same preferences. The percentage bet on the over was found to increase with each point of the total. The total was found to have a positive and significant effect on the percentage bet on the over for both regressions.

For comparison purposes, a total of 51 in the CFL (slightly less than the average of 51.31 for the sample) would be expected to generate 52.85 percent of the money on the over. A likely high-scoring game with a posted total of 60 would expect to generate 59.80 percent on the over. In the AFL, a total of 105 (slightly less than the average of 105.21 for the sample) would expect to generate 55.94 percent on the over. A high total of 120 would expect to generate 60.39 percent on the over.

In neither the sides nor the totals does it appear the sportsbook is attempting to set the price to perfectly balance the book. The sportsbook seems content to attract a higher percentage of bets on the favorite and the over without regard to attempting to eliminate its risk by setting the pointspread or total higher to attempt to even the betting action. The sportsbook even seems less concerned when it comes to road favorites and high totals, as the public overbets these propositions and the sportsbook seems content to let them.

Having an unbalanced sportsbook does not necessarily imply that the sportsbook is pricing to maximize profits, as suggested by Levitt (2004). The sportsbook could be setting prices where the public bets on the side of the proposition which loses more often than it wins, earning profits for the sportsbooks, such as in the NFL (Paul and Weinbach, 2007). It could also be setting prices as a forecast, without regard to the betting percentages of the public, as seems to be the case in the NBA (Paul and Weinbach, 2008).

To attempt to determine if sportsbooks earn profits by pricing at a point other than where the book is balanced, simple betting simulations are presented for both the CFL and the AFL. The first simulations attempt to determine if the samples used in this study exhibit similar betting biases to results found in longer samples for the CFL (Paul and Weinbach, 2007) and the AFL (Borghesi, Paul, and Weinbach, 2009). Biases found in these papers illustrate that bettors of these sports prefer the biggest favorites and the highest totals. Results are shown for the CFL, sides and totals, in tables 5 and 6 and for the AFL, sides and totals, in tables 7 and 8.

Table 5: Simple Pointspread Betting Simulation: Bet the Underdog when Favorite exceeds a certain Pointspread Threshold – CFL (2005-2007)

Favorite Threshold	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
10+	10	14	58.3333%	0.6698	0.3431
7+	28	31	52.5424%	0.1526	0.0006
3+	79	84	51.5337%	0.01534	NA
All	92	106	53.5354%	0.9907	0.1060

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

Table 6: Simple Total Betting Simulation: Bet the Under when Total Exceeds a Certain Threshold – CFL (2005-2007)

Total Threshold	Over Wins	Under Wins	Under Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
55+	13	18	58.0645%	0.8100	0.4039
50+	48	67	58.2609%	3.1536*	1.6043
All	77	102	56.9832%	3.5031*	1.5273

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

Table 7: Simple Pointsread Betting Simulation: Bet the Underdog when Favorite exceeds a certain Pointsread Threshold – AFL (2005-2008)

Favorite Threshold	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
10+	37	53	58.8889%	2.8596*	1.5394
7+	85	105	55.2632%	2.1092	0.6347
3+	200	206	50.7389%	0.0887	NA
All	247	260	51.2821%	0.3334	NA

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

Table 8: Simple Total Betting Simulation: Bet the Under when Total Exceeds a Certain Threshold – AFL (2005-2008)

Total Threshold	Over Wins	Under Wins	Under Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
110+	56	83	59.7122%	5.2781**	3.0210*
105+	118	150	55.9701%	3.8300	1.3892
All	252	265	51.2573%	0.3269	0.2611

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

In the CFL, win percentages on simple strategies of bet the underdog or bet the under do win more than 50 percent of the time, but the only statistically significant results are found for the group of all totals and all totals greater than 50. In the AFL, statistical significance is only found at the highest levels (betting the underdog when there are 10+ point favorites and betting the under when the total is 110+). For the subsample of games with the highest totals, a simple strategy of wagering on the under was found to reject the null of no profitability at the 10 percent level.

The biases in this sample are similar to biases seen in longer samples found in Paul and Weinbach (2007) and in Borghesi, Paul, and Weinbach (2009). The biggest favorites and highest totals are priced slightly too high, with contrarian strategies of betting against these publically popular propositions earning large enough profits to reject the null of a fair bet and, in the case of AFL totals, reject the null of no profitability.

To test if sportsbooks are pricing to maximize profits by exploiting common bettor biases, as suggested by Levitt (2004), we test a few simple betting simulations for the CFL and AFL. Given that bettors tend to prefer favorites and overs (in particular, big favorites and high totals), these simulations test the returns to betting against the most popular betting propositions. Given the availability of the betting percentages, when the game is significantly imbalanced (>70 percent, >60 percent, etc.), we test the returns to a contrarian strategy of betting the underdog or under. Given that the sportsbook is weighted on the favorite and over in these situations, these simulations also represent the return the sportsbook earns by not pricing to balance the book.

Returns to these strategies are presented for the CFL in tables 9 and 10, while returns for the AFL are presented in tables 11 and 12. For each situation, the number of favorite wins, underdogs wins, the underdog win percentage, and log-likelihood ratio tests for a fair bet and no profits are presented. The null of a fair bet implies a win percentage of 50 percent, while the null of no profits implies a win percentage of 52.4 percent, the percentage needed to overcome the commission charged by sportsbooks on bets.

Table 9: Simple Pointsread Betting Simulation: Bet the Opposite of the Public When the Percentage Bet on Favorite Exceeds a Certain Threshold – CFL (2005-2007)

Percentage Bet on the Favorite	Favorite Wins	Underdog Wins	Underdog Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
70%+	18	25	58.1395%	1.1446	0.5753
60%+	40	48	54.5455%	0.7283	0.1657
50%+	61	80	56.7376%	2.5681	1.0778

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

Table 10: Simple Total Betting Simulation: Bet the Opposite of the Public When the Percentage Bet on Under Exceeds a Certain Threshold – CFL (2005-2007)

Percentage Bet on the Over	Over Wins	Under Wins	Under Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
70%+	14	16	53.3333%	0.1334	0.0109
60%+	29	39	57.3529%	1.4759	0.6775
50%+	52	64	55.1724%	1.2436	0.3635

Table 11: Simple Pointspread Betting Simulation: Bet the Opposite of the Public When the Percentage Bet on Favorite Exceeds a Certain Threshold – AFL (2005-2008)

Percentage Bet on the Favorite	Favorite Wins	Underdog Wins	Favorite Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
75%+	55	51	51.8868%	0.1510	NA
70%+	125	99	55.8036%	3.0247*	1.0557
65%+	179	172	50.9972%	0.1396	NA

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

Table 12: Simple Total Betting Simulation: Bet the Opposite of the Public When the Percentage Bet on Under Exceeds a Certain Threshold – AFL (2005-2008)

Percentage Bet on the Over	Over Wins	Under Wins	Over Win Percentage	Log-Likelihood Ratio Test: Fair Bet	Log-Likelihood Ratio Test: No Profits
75%+	78	68	53.4247%	0.6855	0.0639
70%+	129	131	49.6154%	0.0154	NA
65%+	183	182	50.0137%	0.0027	NA

The win percentages are found to be above 50 percent in most of the sample subsets shown in the tables above for the CFL and AFL. These win percentages, however, are not found to be statistically significant against the null of no profitability. In addition, there is only one subset which is found to be statistically significant against the null of a fair bet, the subset of all AFL games where the public bets 70 percent or more on the favorite, which is significant at the 10 percent level.

Overall, there is some slight evidence that sportsbooks may shade the lines and totals slightly to take advantage of known bettor biases, but this pricing by no means earns substantial profits for the sportsbook. In addition, there is little incentive for informed bettors to take contrarian positions in this market as the returns to these simple betting strategies are not statistically profitable.

It appears sportsbooks generally price the CFL and AFL (set pointspreads and totals) as forecasts with a possible slight shading of the line to the big favorites and high totals. This shading of the pointspreads and totals allows the sportsbook to earn slightly higher profits, but are not large enough to invite informed bettors into the fray. It appears that sportsbooks price these relatively small sports mainly as a forecast, not attempting to perfectly balance the book, and earn their commission on losing bets over time, with win percentages on simple strategies expected to hover around 50 percent.

III. Conclusions

The betting markets for the Canadian Football League and the Arena Football league were 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, we tested these hypotheses for the traditional models of sportsbook behavior, where the book is balanced; the Levitt hypothesis, where sportsbooks price to maximize profits; and 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 the CFL and AFL 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, some simple tests were performed on the CFL and AFL data. First, simple betting strategies of betting the underdog and the under were performed. In the CFL, underdogs won slightly more often than favorites, but the results were not found to be statistically significant. In the CFL totals market, for the sample of all totals and the subsample of all totals of 50 or more, the under was found to win often enough to reject the null hypothesis of a fair bet.

In the AFL, big underdogs (10 or more points) were found to win more often than implied by efficiency (nearly 59 percent of the time), while the sample of all underdogs won more than 50 percent of the time, but the results were not statistically significant. In the AFL totals market, the under was found to win nearly 60 percent of the time for games with the highest totals and was found to be significant.

When considering betting percentages and calculating the results when the betting public heavily supports 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 percent 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.

Overall it appears there may be some slight shading of the pointspread and total toward the favorite and the over in the CFL and the AFL. This shading is not great, however, and does not offer much in the way of expected profits for contrarian bettors. 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. This situation results in findings that are more similar to the NBA (Paul and Weinbach, 2008) than the NFL (Paul and Weinbach, 2007).

REFERENCES

- Borghesi, R., Paul, R., and Weinbach, A. (2009) Market Frictions and Overpriced Favorites: Evidence From Arena Football. *Applied Economics Letters*, 16(9), 903-906.
- Golec, J. and Tamarkin, M. (1991). The degree of inefficiency in the football betting market: Statistical tests. *Journal of Financial Economics*, 30, 311-23.
- Gray, P. and Gray S. (1997). Testing market efficiency: Evidence from the NFL sports betting market. *Journal of Finance*, 52, 1725-37.
- 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, R. and Weinbach, A. (2002). Market efficiency and a profitable betting rule: Evidence from totals on professional football. *Journal of Sports Economics*, 3, 256-63.
- Paul, R. and Weinbach, A. (2007). Does Sportsbook.com Set Pointsreads 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, 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.
- 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.