Do Sports Gambling Markets Efficiently Adjust Lines to Incorporate New Information?

Jimmy Torrez, Ph.D
Associate Professor
Department of Finance
College of Business
University of Puerto Rico
Rio Piedras, Campus

Do Sports Gambling Markets Efficiently Adjust Lines to Incorporate New Information?

Gambling markets provide an especially interesting environment under which to study pricing behavior because the fundamental value is always revealed at the close of each contract, unlike asset prices that only provide repeated estimations of the fundamental value. The financial literature has provided a fair amount of work on the efficiency of earnings estimates made by financial analysts as new information becomes available to the market. Likewise, academics have shown a large amount of interest as to the efficiency of gambling markets. There however has been very little work on how gambling markets react to new information. Using methodology similar to that applied to the predictions of financial analysts, this work will evaluate the improvement in college football point spreads as the season progresses and more information on team quality is revealed. College football lends itself to this type of analysis since player turnover insures that at the start of each year there will be a certain amount of uncertainty as to the relative quality of most teams. This paper finds that college football point spreads react in a mildly efficient manner to new information.

Introduction

A necessary aspect of efficient markets is that market participants react quickly to new information. This reaction should be rational and therefore on average correct. This paper borrows from the literature that examines the efficiency of the estimates produced by financial analysts. More specifically it borrows from the literature that estimates how analysts correct their estimates as new information becomes available to the market. College football lends itself to this type analysis since college football players only have 4 years of eligibility to play, which insures a certain amount of uncertainty as to the relative quality of each team. If this market is indeed efficient, the accuracy of point spreads[1] should improve as more information becomes available after each week's performance. In addition, the amount of games each week in college football is about 4 times that of National Football League (NFL). It is for these reasons that college football provides a better platform to evaluate how changes in information affect point spreads than does professional football.

Using data from the college football point spreads this paper demonstrates a small game-to-game improvement in the predictive capabilities of the lines bookmakers set each week. This is tempered by the fact that this small improvement due to new information is not large enough to offset other factors that make it more difficult to predict the outcome of games later in the season than at the beginning of the season.

[1] In a point spread wager the payoff depends on the difference in points scored by the two opposing teams.

Literature Review

The characteristics of efficient markets do not allow participants to earn a persistent risk adjusted profit from new information (Fama 1971). Even before the seminal work by Fama, researchers were interested in the efficiency of financial markets (Roberts 1959). The speed at which asset prices adjust have also been studied for quite some time (Patell and Wolfson 1984).

Scholars also have been interested in how financial analysts react to new information. Since financial analysts compete with each other and sell their forecasts to market participants, one would not expect them to ‘‘leave money on the table’’ by inefficiently using all this publicly available information (Keane and Runkle, 1998). Keane and Runkle do however find inefficiency in the earnings estimates of these analysts. Basu and Markov (2004) on the other hand demonstrate that analysts actually generate efficient estimates when Least Absolute Deviation regressions are used instead of Ordinary Least Square regressions.

A similar line of research concentrates on how analysts react to new information (DeBondt and Thaler (1986), Brous (1992), Abarbanell and Bernard (1992) and Easterwood and Nutt (1999)). Although there is some disagreement whether earnings estimates overreact, underreact or are overly optimistic, the above authors agree that analysts do not react efficiently to new information.

Although most work on efficient market hypothesis (EMH) has concentrated on financial markets there is a growing literature that applies EMH to other types of markets. One such application is to gambling markets (Pankoff (1968), Sauer (1998) and Paton, Siegel and Williams (2009)). These markets provide an especially interesting environment under which to study pricing behavior because the fundamental value is always revealed at the close of each contract, unlike asset prices that only provide repeated estimations of fundamental value.

Similar to evidence on earnings forecasts the majority of evidence suggests that gambling markets are inefficient. There are other similarities. Efficient Market Hypothesis is plagued with a number of abnormalities when it comes to explaining the behavior of asset prices. Examples of this are the Equity Premium puzzle, the Closed-End Fund puzzle, the Dividend puzzle, the January effect and the Small Stock effect among others. In sports betting there exist a favorite bias and a visiting team bias (Gray and Gray (1997) and Paton, Siegel and Williams (2009)). Therefore a bettor using a strategy of only betting on the underdog or the home team will win more that 50% of his bets. This implies there are a number of "irrational" bettors who are betting on the favorite and/or the home team even thought the spread is on average biased against these bets.

Until recently many, if not most researchers believed line-makers acted only as intermiedaries and did not take a position on the outcome of the game. Pankoff (1968) for instance states "As a rule, the book-maker's role is primarily that of broker. His income is derived from keeping a percentage of all winning bets." Similarly, Lee and Smith (2002) write, "Bookies do not want their profits to depend on the outcome of the game. Their objective is to set the point spread to equalize the number of dollars wagered on each team and to set the total line to equalize the number of dollars wagered over and under. If they achieve this objective, then the losers pay the winners $10 and pay the bookmaker $1, no matter how the game turns out." Levitt (2004) and Strumpf (2003) have provided evidence that bookmakers do indeed take stakes in the outcome of sporting contests. Paul, Weinbach and Paul (forthcoming) on the other hand show that line movements seem to take advantage uninformed betters.

Whether line-makers take advantage of uniformed betters or whether informed bettors move the spread is irrelevant for this work. The purpose of this work is to evaluate whether the accuracy of betting markets improves as the season progresses. As stated, college football lends itself to this type analysis since college football player only have 4 years of eligibility to play. This turnover insures that at the start of each season there will be a certain amount of uncertainty as to the relative quality of the teams. In addition "young" teams have an opportunity to improve throughout the season. Therefore there should be room for improvement in the accuracy of spreads as the season progresses. If this market is indeed efficient the accuracy of the point spread should improve as more information becomes available after each week's performance. Bowl Series Championship colleges (formally division I-A) consists of between 117 to 120 teams each year being analyzed, this increases the amount of games played weekly relative to the NFL which only consists of 32 teams. College football therefore provides a better platform to evaluate how changes in information affect point spreads than does professional football.

Data

The data used in this paper comes from 3 different sources. Data for 2001 and 2002 will be from the 2003 *Gold Sheet, College and Pro Football Annual*. Data for 2003 to 2007 will be from the 2008

Phil Steele's College Football Preview Magazine and data form 2008 and 2009 will be from *Phil Steele's College Football Preview Magazine* 2010 edition.

Methodology

The basic methodology in this work will be based on the basic regression used by Abarbanell and Bernard (1992) who examine how stock analysts incorporate new information in their earnings estimates. They run the following regression.

(1) $(E_t - F^{t-1}_t)/P = \alpha_0 + \alpha_1 \, PERF_{t-1}$

Where

F^{t-1}_t is forecasted earnings for period t made at period t-1.
E_{t-i} is actual earnings at period t-i.
P is the price of the stock at the time the forecast was made.
$PERF_{t-1}$ is the prior period earnings change [$PERF_{t-1} = (E_{t-1} - E_{t-2})/P$]

In this simple regression they find that α_1 is positive implying that analysts underreact to prior period changes in earnings[2].

Before continuing it may be prudent to briefly explain how point spreads work. Point spreads (PS) are typically reported as the number of points by which one team is "favored" to beat another team. Since teams will be evaluated both separately and as a whole the methodology to account for the difference in points (DP) will be different from much of the point spread literature. In this case the actual difference in points is defined as the points scored by the opponent of the team being examined less those scored by that team. When the team in question

[2] Easterwood and Nutt (1999) do a more in-depth analysis and find that Financial analysts are actually overly optimistic, overreacting to positive earnings changes and underreacting to negative earnings changes.

is favored a bet on that team pays off when *DP* - *PS* < 0, bets on underdogs pay off when *DP* - *PS* >0, and all bets are refunded when *DP* = *PS*.

Unfortunately point spreads only come out the week of the contest so it is impossible to observe a change in predicted point differentials due to relative performance changes in prior periods. In addition, point spreads depend greatly on the perceived relative quality of the two opponents. For this reason the exact methodology of Abarbanell and Bernard cannot be used. Fortunately unlike earnings which tend to be autocorrelated from period to period, one should not expect the difference between the actual margin of victory or loss and the point spread to be correlated from week to week. If this market is efficient however one should expect a week-to-week relationship between actual point differentials and those predicted by the point spread to improve as more games are played. Therefore equation (1) above will have to be modified in the following manner.

(2) $(DP_t - PS_t) = \alpha_0 + \alpha_i (DP_{t-i} - PS_{t-i})$

Where

DP is the actual point differential between the opponent and the team in question.
PS is the point differential predicted by the point spread.
i is the lag

In general one would hypothesize that the coefficients of the actual point differentials vs that predicted by the point spread to be negative. A positive coefficient would imply that predictions become worse as the season progresses, which is inconsistent with EMH. Alternatively a coefficient equal to zero would indicate that there is no change in the accuracy of point spread estimates from week to week.

Since it is likely that most teams have a significant amount of turnover and that information improves throughout the season the first game of the season will be omitted. In addition, for every favorite with a negative point spread there exists an underdog with a positive point spread. Therefore simply taking and average will be biased since the favorites and underdogs will cancel each other out. Therefore averages will be taken for only the favorites[3] and the absolute value of the actual difference in points minus the point spread (DP-PS).

Results

Table 1 below demonstrates the basic statistics of (DP-PS) for favorites using both the raw data and the absolute value. It also contains absolute values for all teams in the dataset.

Table 1

Basic Statistics (DP-PS)

	Average	Median	Standard Deviation	Observations
Favorites	-0.361	0.000	15.516	6171
Absolute Value of Favorites	12.192	10.000	9.603	6171
Absolute Value of all teams	12.207	10.000	9.610	12101

As expected the average and median is much smaller while the standard deviation is much bigger for the favorites alone compared to the absolute value of the entire data set. The reason for this is that two teams have the same (DP-PS) in under the absolute value measure, while for the

[3] Nearly all favorites in the data set have an offsetting underdog. The exceptions are those games in which the team in question plays and non-BCS eligible team. In all but one of those cases the team in question is favored.

favorites it will only be counted once. As can be seen from the median value, favorites actually cover the point spread 50% of the time implying bookmarker on average set the point spread correctly, implying some form of efficiency in this market. Another interesting piece of information we can deduce from these basic statistics in examining the median of the absolute value measure is that 50% of the time the favorite is expected to win by 10 points or more. The standard deviation is smaller for the absolute value since only one side of the athletic contest is counted compared to the favorite to where it is possible that the favorite covers the spread [(DP-PS)<0] or not [(DP-PS)>0]. This of course increases the possible range of the statistics which adds to the standard deviation.

In order to examine whether there exists evidence that bookmakers predictions improve with new information throughout the season the regression in equation 2 is preformed. This regression is run with the absolute value of (DP-PS) because of the difficulty in interpreting the coefficients with the raw data. Using only favorites or underdogs would not work either since it is common for a team to be a favorite one week and an underdog the next. The regression is first performed with the entire dataset using dummy variables for each year and team. The results are shown in Table 2 below.

Table 2

Regression result [entire dataset]*

$$(DP_t\text{-}PS_t) = \alpha_0 + \alpha_i (DP_{t\text{-}i}\text{-}PS_{t\text{-}i})$$

Intercept	Coefficient	Year Dummies	Team Dummies
10.673 (0.961)*	-0.019 (0.009)*	Yes	Yes
	R-squared	Adjusted R-squared	
	0.018	0.007	

* Standard Errors using the Newey-West correction for autocorrelation and heteroscedasticity are in parenthesis

Since it is likely that certain teams are followed more closely than others and provide more information to the market this regression is also run for each team separately using year dummies. This has advantages to just using dummy variables since we can estimate the amount of teams for which spreads improve throughout the seasons. Of the 120 teams for which regressions are run, 16 have positive coefficients and 104 have negative coefficients. Of those with a negative coefficients, 40 are significant at least at the 10% level, none of the positive coefficients are significant. So there seem to be evidence that point spread slightly improve their predictive effects from week to week for a significant amount of teams. The result are shown in table 3 below.

Table 3

	# of teams negative coefficients on the (DP-PS) variable	# of teams negative coefficients on the (DP-PS) variable
All	104	16
Teams who’s coefficients are also significant	40	0

Given the results above it would seem reasonable that (DP-PS) should be smaller during the later part of the season than at the beginning of the season. The average for the first 4 games of the season for which there exists a point spreads are compared to the last 4 games of the season. Given that point spreads do not exist for all games, when 4 point spreads do not exist during the first or last 5 games of the season the average is taken with 3 games. Of the 120 teams only 56 have smaller absolute (DP-PS) during the later part of the season, of those only 5 were significant using t-test at the 10% level. Of the 64 teams who's average absolute (DP-PS) was bigger during the first part of the season 5 were also significantly different at the 10% level.

Another possibility is that variance improves throughout the season. This is however not the case. Of the 120 teams only 48 have smaller variances the last 4 games of the season than the first, of those only 6 have significantly smaller variances using the Folder F test at the 10% level. Twenty-three of the remaining 72 teams have positive and significantly larger variances for the last 4 games of the year compared with the first 4.

Table 4

	# of teams with smaller average (DP-PS) the last 4 games of the season than the first 4 games	# of teams with larger average (DP-PS) the last 4 games of the season than the first 4 games	# of teams with smaller variance for (DP-PS) the last 4 games of the season than the first 4 games	# of teams with larger variance for (DP-PS) the last 4 games of the season than the first 4 games
All	56	64	48	72
Significantly different Averages or Variances	5	5	6	29

The results above can be explained if there is something during the latter part of the season that makes it more difficult to predict the outcome. Some possible explanations may be rivalry games and injuries. Rivalry games tend to be played with more emotion and therefore it may be more difficult to predict the outcome of these games. It just so happens that there tend to be many more rivalry games during the end of the football season. Injuries tend to be cumulative in that injuries towards the beginning of the season may last throughout the season. It is more difficult to evaluate how injured player will perform as well as how backup players will perform that substitute injured players who cannot play.

The above results are not surprising given that the improvement in the week-to-week coefficient is so small relative to most point spreads. It is probable that factors more common toward the end of the season make it difficult to predict the impact they will have on the final score. This will of course increase the variance of (DP-PS) relative to earlier games.

Conclusion

It is expected that if markets are efficient market participants will react quickly to new information. Efficient market theory predicts that on average this reaction to new information will be correct. This paper borrows from the literature that examines the efficiency of estimates of financial analysts. More specifically it borrows from the literature that estimates how analysts correct their estimates as new information hits the market.

College football lends itself to this type analysis since college football player have a certain amount of turnover that insures that at the start of each year there will be a certain amount of uncertainty as to the relative quality of the teams. Using data from college football point spreads this paper demonstrates a small but statistically significant game-to-game improvement in the predictive capabilities of the lines bookmakers set. Unfortunately this small improvement is not large enough to offset other factors that make it more difficult to predict the outcome of games later in the season than at the beginning of the season.

Abarbanell, Jeffery S., and Victor L. Bernard, (1992), Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior," Journal of Finance 47, 1181-1207.

Basu, Sudipta and Stanimir Markov "Loss Function Assumptions in Rational Expectations Tests on Financial analysts'Earnings Forecasts" Journal of Accounting and Economics 38 (2004) 171–203.

Brous, Peter A. 1992, Common Stock offerings and earnings expectations: A test of the release of unfavorable information, Journal of Finance 47, 1517-1536.
DeBondt, Werner F.M., and Richard M. Thaler, 1990, "Do security analysts overreact?," American Economic Review 80, 52-57.

Easterwood, John C. and Stacey R. Nutt " Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism?," *The Journal of Finance*, Vol. 54, No. 5 (Oct., 1999), pp. 1777-1797

Gray, P. and Gray, S. (1997). 'Testing market efficiency: evidence from the NFL sports betting market', *Journal of Finance*, vol. 52, (September), pp. 1725–37.

Keane, M.P., Runkle, D.E., 1998. Are financial analysts' forecasts of corporate profits rational? *Journal of Political Economy* 106 (4), 768–805.

Lee, Marcus, and Gary Smith (2002) *Journal of Behavioral Decision Making*, Vol. 15 Issue 4**,** August pp. 329 - 342

Levitt, Steven. (2004). "Why Are Gambling Markets Organized So Differently from Financial Markets?".*Economics Journal*, Vol. 114 No. 495:223–46.

Pankoff, Lyn D. (1968) "Market Efficiency and Football Betting," *The Journal of Business*, Vol. 41, No. 2 April, pp. 203-214

Patell J.M., Wolfson M.A. (1984) "The intraday speed of adjustment of stock prices to earnings and dividend announcements" *Journal of Financial Economics*, 13 (2), pp. 223-252.

Paton, Davd Donald S. Siegel and Leighton Vaughan Williams, "The Growth of Gambling and Prediction Markets: Economic and Financial Implications," *Economica* (2009) 76, 219–224

Paul, Rodney, Andrew Weinbach and Kristin Paul (forthcoming) "Noise Trading and Pointspread Movements in the NCAA Football Betting Market" *Journal of Business, Industry and Economics*

Peel, David A. and Law, David A., (2009) A More General Non-Expected Utility Model as an Explanation of Gambling Outcomes for Individuals and Markets. Economica, Vol. 76, No. 302, pp. 251-263, April.

Phil Steele's College Football Preview Magazine (2008)

Phil Steele's College Football Preview Magazine (2010)

Roberts, Harry V. (1959) "Stock Market 'Patterns" and Financial Analysis: Methodological Suggestions." Journal of Finance, 14 (March), 1-10.

Sauer, Raymond D.,(1998) "The Economics of Wagering Markets," *Journal of Economic Literature, Vol. XXXVI December, pp. 2021–2064*

Strumpf, Koleman. (2003). 'Illegal sports bookmakers', *Mimeo,* University of North Carolina. Department of Economic.