Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games

Author(s): Richard A. Zuber, John M. Gandar, Benny D. Bowers

Source: *The Journal of Political Economy*, Vol. 93, No. 4 (Aug., 1985), pp. 800-806
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/1832139

Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games

Richard A. Zuber, John M. Gandar, and Benny D. Bowers

University of North Carolina at Charlotte

In this paper we test the efficiency of the gambling market for National Football League games. Two efficiency tests are conducted. The first test is derived from the finance literature on market efficiency, while the second test is based on a market's being efficient when the rate of return on any gambling strategy based on publicly available information approximates the bookmaker's commission. While the first test is found to be too weak to establish conclusions about the efficiency of the NFL gambling market, the second test results, showing the existence of profitable betting opportunities, indicate that speculative inefficiencies exist in this market.

I. Introduction

In recent years a number of studies, employing tests of efficiency developed for analyzing securities markets, have examined the efficiency of racetrack gambling. Results are mixed: certain studies have found these markets to be efficient (Snyder 1978; Ali 1979; Figlewski 1979); others have found inefficiencies that allow for profitable wagering (Hausch, Ziemba, and Rubinstein 1981; Asch, Malkiel, and Quandt 1984). In this paper we extend the examination of gambling market efficiency to the gambling market for National Football League (NFL) games. The only previous work in this area,

We would like to thank Jim Marsden, Paul E. Smith, and an anonymous referee for valuable comments.

[*Journal of Political Economy*, 1985, vol. 93, no. 4]

the study of Vergin and Scriabin (1978), investigated biases in the setting of point spreads for NFL games and developed betting strategies based on these biases that allowed for profitable wagering. The intent of the present paper is to test for efficiency in this gambling market more directly than do Vergin and Scriabin. We conduct both a "weak" test of efficiency, the ability of the Las Vegas gambling "line" to predict point spreads, and a stronger test, the ability of an explanatory model of actual point spreads to predict game outcomes and earn speculative profits.

II. Initial Efficiency Tests

The conventional test of market efficiency is often presented in the form of the hypothesis that forward prices are the best unbiased forecast of future spot prices.[1] The direct equivalent of this efficiency test for the NFL gambling market is to estimate the equation

$$PS_{it} = b_0 + b_1 VL_{it} + u_{it}, \tag{1}$$

where PS_{it} is the actual point spread on the ith game in the tth week, VL_{it} is the Las Vegas gambling line,[2] and u_{it} is the error term. It should be understood that VL_{it} is not the bookmaker's prediction of the game's outcome. While the initial line is based on an expert's opinion of the game's outcome, thereafter the bookmaker adjusts the line as gambles are made in order to avoid exposure to unnecessary risk. Thus the line on a game moves to reflect the collective judgment of gamblers about its outcome. If gamblers are using the available information efficiently, then we might expect the final point spread to be the best unbiased forecast of the game's outcome. In terms of equation (1) the efficient markets test is the null hypothesis that jointly $b_0 = 0$ and $b_1 = 1$.

To obtain F-values for this test, equation (1) was estimated by ordinary least squares (OLS) for each week of the 1983 NFL regular season. For 13 of the 16 weeks the null hypothesis cannot be rejected at conventional levels of significance. However, before concluding that efficiency prevails, consider the alternative hypothesis that the Las Vegas line is unrelated to the actual game point spread. In terms of equation (1) this is the joint test that $b_0 = b_1 = 0$. This alternative hypothesis cannot be rejected for 15 of the 16 weeks. That is, the

[1] For a convenient summary of this and other tests see de Leeuw and McKelvey (1984).

[2] The Las Vegas line used throughout the paper was the last one publicly available, the line published in the media on Fridays over the season. For consistency both PS_{it} and VL_{it} are defined from the home team's perspective. In this paper when $VL_{it} > 0$ ($<$ 0), the home team is the favorite (underdog).

extreme alternative hypothesis is as consistent with the sample data as is the efficiency hypothesis.[3] An alternative testing strategy is required.

To initiate this, consider the following definition of efficiency in a gambling market (see Asch et al. 1984). Such a market is efficient when the rate of return to any gambling strategy based on publicly available information approximates the bookmaker's *vigorish* (or track take). For point spread gambling this commission is approximately 5 percent. That is, in an efficient football gambling market gamblers pursuing any strategy should earn a negative rate of return close to 5 percent. While efficiency in this sense is difficult to establish, it is possible to examine whether there exists sufficient divergence from efficiency in the gambling market to permit profitable gambling strategies. If such gambling opportunities exist we could conclude that inefficiencies exist. The remainder of this paper is concerned with constructing this alternative test of inefficiency for NFL gambling.

III. Explanatory Model Specification and Results

In this section we develop a plausible explanatory model systematically utilizing publicly available information on NFL games. We see this as a necessary first step for the prediction of game point spreads required for the gambling simulations carried out in the next section.

The actual result of any football game, as measured by the point spread, can be viewed as the net outcome of the simultaneous efforts of the two teams involved. Thus the game point spread is viewed here as the dependent variable in a reduced-form equation, while the independent variables are the net efforts of the two teams. These net efforts, in turn, can be represented by the game statistics and the teams' characteristics. More formally, let PS_{it} be the point spread (measured as the home team's score minus the visiting team's score) of the ith game in the tth week. Then the argument above sees PS_{it} as the dependent variable in the equation

$$PS_{it} = \mathbf{B}'X_{it} + u_{it}, \tag{2}$$

where $\mathbf{B}'$ is a vector of coefficients to be estimated, X_{it} a matrix of observable variables, and u_{it} an error term. Each of the game and team variables in the matrix X_{it} is defined in the same fashion as PS_{it}; that is, each consists of the first difference between the value of the variable for the home team and its value for the visiting team. The explanatory variables used in the estimation of equation (2) for this paper include yards rushed (YR_{it}), yards passed (YP_{it}), number of

[3] Complete results of these tests are available on request from the authors.

TABLE 1

OLS ESTIMATES OF EQUATION (2): 1983 REGULAR NFL SEASON

	First 8 Weeks	Entire 16 Weeks
Constant	1.547	1.625
	(.581)**	(.514)**
YR	.047	.041
	(.010)**	(.009)**
YP	.044	.045
	(.007)**	(.006)**
W	.697	.574
	(.300)*	(.216)**
F	−2.299	−2.036
	(.418)**	(.362)**
IN	−2.619	−2.788
	(.344)**	(.308)**
PN	−.424	−.309
	(.171)*	(.146)*
PP	−.217	−.244
	(.050)*	(.046)**
R	−.319	−.309
	(.130)*	(.117)**
R^2	.733	.727
F	37.28	71.63
Df	103	215

NOTE.—Standard errors are in parentheses.
* Significant at the 5 percent level.
** Significant at the 1 percent level.

wins previous to the game (W_{it}), fumbles (F_{it}), interceptions (IN_{it}), number of penalties (PN_{it}), proportion of passing plays attempted to total offensive plays (PP_{it}), and number of rookies (R).[4] A priori expectations are for positive signs on the first three explanatory variables and negative signs on the remainder. Data for all of these variables for the 1983 season were gathered entirely from publicly available information.[5]

While equation (2) was estimated by OLS for a number of subdivisions of the 1983 season, because of the similarity of results table 1 presents only those for the first half of the season as well as the entire season. The results are gratifyingly neat and clear. All of the explanatory variables are significant and have their expected sign. As measured by the R^2 value, this simple model explains nearly three-fourths

[4] These explanatory variables were found significant in all models examined over all subdivisions of the 1983 season. Of them perhaps only PP_{it} requires elucidation: this variable was introduced into the model to capture the added risks of the passing-based offenses frequently employed by trailing teams in their efforts to catch up.

[5] The data set used in this study was gathered from the *Sporting News*, the *Charlotte Observer*, the *1983 NFL Media Information Book*, and individual 1983 team media guides.

of the squared variation in actual point spreads. Finally, estimation of equation (2) sequentially over the last 8 weeks of the season shows that conventional F-tests are unable to reject the null hypothesis of parameter stability. That is, our explanatory model can be utilized for predicting point spreads in the latter part of the season.

IV. Prediction and Gambling Simulation

Predicted point spreads (PPS_{it}) for the second half of the 1983 NFL regular season ($t = 9, \ldots, 16$) were obtained using the form of the explanatory model developed above. This equation was estimated using the data for week 1 to week $t - 1$. The resulting parameter estimates were then applied to team-specific averages (similarly computed through week $t - 1$) of the relevant home and visiting team explanatory variables. The point spread predictions for the tth week thus represent ex ante predictions utilizing only information available prior to the games in that week.

The predicted point spreads were then used in a number of gambling simulations. In these simulations a bet was made on a game if $|PPS_{it} - VL_{it}| \geqq \lambda$, where λ took on the values of 0.5, 1, 2, and 3. Conditional on a bet's being made, the team to bet on is determined by the sign of ($PPS_{it} - VL_{it}$): if this expression is positive, the gamble is made on the home team; if negative, the gamble is made on the visiting team. Finally, the bet is won if the team gambled on beats the spread. This requires that the sign of ($PS_{it} - VL_{it}$) coincide with the sign of ($PPS_{it} - VL_{it}$).[6]

The results of the simulation using $\lambda = 0.5$ are summarized in table 2. With this filter 102 bets were placed over the 8 weeks of simulated gambling, 60 of which won for a winning percentage of 59 percent. This occurred despite 2 comparatively poor weeks, weeks 12 and 13, when the winning percentage fell to 46 and 38 percent, respectively. While table 2 contains only the results for $\lambda = 0.5$, it should be noted that all four strategies produced winning percentages greater than the 52.4 percent needed to break even.[7]

To further illustrate the potential profitability of the simulation for $\lambda = 0.5$ assume \$110 is bet on each game in the ninth week that meets the criterion of $|PPS_{it} - VL_{it}| \geqq 0.5$. Since all 14 games meet this

[6] The VL_{it} that determines whether a bet is won or lost is the line at the time the bet is made. This represents a major difference between football and racetrack gambling.

[7] Gambling on an NFL game is carried out on the "11 for 10" rule: the gambler must lay out \$11 for every \$10 he or she wishes to win. The percentage of winning bets (WP) necessary to break even, 52.4 percent, is obtained by setting the expected value of the random variable, a gamble $WP(10) + (1 - WP)(-11)$, equal to zero.

TABLE 2

GAMBLING SIMULATIONS (λ = 0.5), FINAL 8 WEEKS OF 1983 SEASON

Week	Number of Bets	Number of Wins	Wins/Bets (%)	Sum Gambled ($)	Net Return ($)	Rate of Return (%)	Final Wealth ($)
9	14	8	57.1	1,540	140	9.1	1,680
10	13	8	61.5	1,430	250	17.5	1,930
11	13	8	61.5	1,430	250	17.5	2,180
12	13	6	46.2	1,430	−170	−11.5	2,010
13	13	5	38.5	1,430	−380	−26.6	1,630
14	12	8	66.7	1,320	360	27.3	1,990
15	11	10	90.9	1,210	890	73.6	2,880
16	13	7	53.8	1,430	40	2.8	2,920

criterion in week 9, $1,540 is laid out in bets. The net return in this week is $140 (the eight winning bets produce $800 in winnings; the six losing bets cost the gambler $660). Continuing the process for the remaining 7 weeks of the season results in a final wealth position of $2,920 representing net winnings of $1,380.

One important question remains: Are these results true exploitations of market inefficiencies, or could they be obtained simply by chance? The probability of randomly selecting a winning gamble is .5. When the normal approximation to the binomial distribution is used, the probability of randomly selecting 60 winning bets in 102 gambles is less than 5 percent. Thus it is reasonable to suppose that the profitable gambling simulations developed above represent true exploitations of inefficiencies in the NFL gambling market.

V. Conclusions

The purpose of this paper was to investigate the speculative efficiency of the gambling market for NFL games. Since we can show that profitable gambling opportunities exist, speculative inefficiencies appear to be present in this market. These results are consistent with the findings of Vergin and Scriabin (1978) for the 1969–74 NFL seasons.

Finally, several points of caution are warranted. First, our study was conducted over a single NFL season. Simulations over additional seasons are required before confidence in these results can be assured. Second, there is a considerable difference between the demonstration that our predictive mechanism could have made money in the past and its ability to produce profitable returns in the future. Third, while our results suggest that speculative inefficiencies exist in this market, this does not necessarily imply market inefficiency. There are considerable information costs involved in assembling and processing the required data. Also, because gambling is illegal in most states, there may be large transactions costs involved.

References

Ali, Mukhtar M. "Some Evidence of the Efficiency of a Speculative Market." *Econometrica* 47 (March 1979): 387–92.

Asch, Peter; Malkiel, Burton G.; and Quandt, Richard E. "Market Efficiency in Racetrack Betting." *J. Bus.* 57 (April 1984): 165–75.

de Leeuw, Frank, and McKelvey, Michael J. "Price Expectations of Business Firms: Bias in the Short and Long Run." *A.E.R.* 74 (March 1984): 99–110.

Figlewski, Stephen. "Subjective Information and Market Efficiency in a Betting Market." *J.P.E.* 87 (February 1979): 75–88.

Hausch, Donald B.; Ziemba, William T.; and Rubinstein, Mark. "Efficiency of the Market for Racetrack Betting." *Management Sci.* 27 (December 1981): 1435–52.

Snyder, Wayne W. "Horse Racing: Testing the Efficient Markets Model." *J. Finance* 33 (September 1978): 1109–18.

Vergin, Roger C., and Scriabin, Michael. "Winning Strategies for Wagering on National Football League Games." *Management Sci.* 24 (April 1978): 809–18.