2014

Testing the Efficiency of the NFL Point Spread Betting Market

Charles L. Spinosa
Claremont McKenna College

CLAREMONT MCKENNA COLLEGE

Testing the Efficiency of the NFL Point Spread Betting Market

SUBMITTED TO

Professor Serkan Ozbeklik

AND

DEAN NICHOLAS WARNER

BY

Charles L. Spinosa

FOR

SENIOR THESIS

Fall 2014
December 1, 2014

I. Abstract

This paper examines the efficiency of pricing in the NFL point spread betting market, as hypothesized by the Efficient Market Hypothesis, through both statistical and economic tests. This market provides a simpler framework to test such economic hypotheses than conventional financial markets. Using a larger sample size than past literature, this paper finds that while the market is efficient in the aggregate sense, there are still some profit opportunities which imply pricing inefficiencies.

I. Contents

I. Introduction

The Efficient Market Hypothesis (EMH), first formalized in 1970 by Eugene Fama, states that financial markets are "informationally efficient". The primary implication of this hypothesis is that in the face of uncertainty the price of a good serves as an unbiased predictor of the good's future value given all publically available information. In inefficient markets, in contrast, there are better alternative ways to gauge the value of a good, and this asymmetry can be exploited for profits. Empirically testing the EMH, however, is a difficult task; because of the arbitrarily long time between the purchase of a good at its original price and the realization of its value at sale it is hard to determine if the price of the good was accurate. Beginning with Lyn D. Pankoff in 1968 economists identified the NFL point spread betting market as an analogue to financial markets for the purposes of testing market efficiency. Subsequent research has branched into other point spread betting markets as well. (Woodland 1994, Woodland 2001, Gandar 2001).

While unintuitive, the NFL spreads market has some important similarities to financial markets that enable this comparison. In both markets the information regarding past performance and price (spread) of a good (bet) is publicly available, the final value of the good is unknown at the time of purchase, and many self-interested parties are in competition. The NFL market however, in contrast to financial markets, has a defined end period, after which the true value of the good is realized to everyone. This finality allows for much more straightforward testing of the EMH.

Previous research has produced mixed results regarding the market's efficiency, with inefficiencies being relatively minor and transient (Sauer 1998). This paper expands on past research on the efficiency of the NFL spreads market by using a larger sample size than most, twenty one years of regular season games, in order to better assess the existence of profitable market inefficiencies, even if price is an efficient indicator of value in aggregate. As expected given past literature, the NFL point spreads betting market between 1992 and 2012 is efficient in the aggregate sense; Las Vegas point spreads are an unbiased, minimum variance estimator of final point differences. Despite this general market efficiency, however, certain betting strategies were found to be profitable over this time period, indicating that certain market inefficiencies do exist. In the next section an overview of the NFL point spread betting market is provided. Section II contains an overview of the structure of the NFL point spread betting market, Section III contains a review of past literature, Sections IV-VI contain methodology, data, and results. Section VII summarizes the paper.

II. The NFL Point Spread Betting Market

The most common bet placed on NFL games is a straight bet. In this betting format, a bettor chooses to bet on a team based on a published handicap called the "spread" or "line". Suppose team **A** is playing team **B** and the line has team **A** favored by five points. Traditionally, this line will be published as "Team **A** -5"[1]. A bet on team **A** will win (also called "covering the spread") if team **A** wins the game by six or more points. The opposite bet, "Team **B** +5", will win ("cover the spread") if either team **B** wins the game, or if team **B** loses by four or fewer points. In the event that team **A** wins

[1] This definition of the line, being expressed from the perspective of the favorite team, is the most common format published by odds makers. Lines can also be expressed relative to the home team.

by exactly five points, the bet is a "push" and the money bet at this line is refunded[2]. The payouts for this type of straight bet operate under the "eleven for ten" rule; in order to win \$10 a bettor must risk \$11[3]. This 10% commission is called the *vigorish* and serves to compensate the book maker. In order for a bettor to break even with betting unit x and *vigorish* v the bettor must win with probability p such that:

$$p \times x = (1 - p) \times ((1 + v) \times x) \qquad (1)$$

When v is 10%, as is commonplace, this equation can be solved to find that the bettor must pick winning bets 52.38 percent of the time. Similarly, given a winning percentage of p, a bettor's profits π per unit bet x can be calculated using the following formula:

$$\pi = (p \times x) - ((1 - p) \times (1 + v) \times x) \qquad (2)$$

While the above example explains how a single bet works, the point spread betting market as a whole has a much longer life cycle. Traditionally, a small group of line makers in Las Vegas will come to a consensus early in the week and publish an "opening" spread or line. This opening spread is then adopted by (virtually) every sports book such that the different books all open with a consistent line. Traditional wisdom says that this opening line is chosen in order to draw an even volume of bets on each side of the line by the time the book is closed to additional bets[4]. If this is achieved then the

---

[2] It is worth noting that spreads can be set to half points, such as "Team **A** -3.5" in order to avoid pushes.
[3] In the past, and possibly in modern day illegal gambling, the *vigorish* collected by book makers went as high as 20%, and "push" bets were collected by the book keeper instead of refunded.
[4] Levitt (2004) and Paul et al. (2011) suggest that book makers may not choose spreads to equalize betting but instead to maximize book profits. The implications of this new research with regards to betting strategies will be addressed later in the paper. For the

book maker will risklessly profit from the *vigorish* collected on losing bets, while covering all of the winning bets with revenues from losing bets.

In this sense the opening line is not chosen to be a predictor of the final game point differential. If line makers predict that the public will bet irrationally then they can select a spread in such a way that despite the public's biases, the total volume of bets will be roughly equal. For example, if line makers predict that bettors will overly bet on a favored New York team (due to extreme loyalty), they could set the opening line at "New York -5" even though the truly unbiased prediction of the outcome is "New York -4". By effectively raising the price of betting on New York (by making the criteria for payoff more difficult), line makers counteract the biased public, resulting in more even books. On the other hand, line makers must also be wary of sophisticated bettors who are likely to take advantage of these line biases. The opening line should then reflect the relative weights that line makers give the uninformed public bettor versus the expert bettor.

After the opening line is posted betting begins. As bets flow in the spread may change in order to incentivize bettors to bet one way or another. If, for example, a large number of bets come in for "Team **A** -5" the book maker may change the spread for future bets to "Team **A** -6". At this new betting line team **A** will have to win by an even greater margin in order for a bet on team **A** to pay off, effectively increasing the "price" of a bet on **A** and decreasing the "price" of a bet on **B** at the new spread. This change in

---

purposes of introducing the point spread betting market the traditional view of book makers as unbiased brokers will be used.

price should cause more bets to come in for team **B**, balancing the book[5]. Additionally, new information, such as injury reports or weather forecasts, may arise during the week which causes the spread to shift. Regardless of why shifts in the spread might happen, one key aspect of point spread betting is that once a bet is made at a certain spread the bet is locked in at that spread. This contrasts with some other betting styles such as the Pari-Mutuel system used in horse racing[6]. At the end of the betting period, usually minutes before the game starts, sports books stop taking bets. The final spread at the time of closing, and the one used in this paper, is usually called the closing line or public line. The closing line should, in aggregate, reflect the public's betting preferences.

III. Literature Review

The past literature regarding the efficiency of the NFL point spread betting market is very mixed. Many authors claim to have found either statistical evidence of market efficiency, or economic evidence of inefficiency (in the form of profitable betting patterns). The first meaningful work analyzing the NFL point spread betting market was done by Pankoff (1968) in an attempt to test market efficiency. Pankoff tested the

---

[5] One potential outcome of adjusting spreads in an attempt to balance books is that the book maker can get "middled". Continuing the example above, if the betting public did not bet at the new spread of "Team **A** -6", the book maker might adjust the spread again to "Team **A** -7". If this newest spread finally attracts bettors for team **B** the book maker is exposed to a risky situation if team **A** wins by 6. In this case, team **A** covers the spread at the original line of -5, while team **B** covers the spread at the new line of +7. The book keeper then has to pay out bets to those on each side. Superbowl XIII is infamous due to a large number of books getting "middled" and losing large sums of money on the game. Most books now are extremely hesitant to move more than three or four points for fear of this phenomenon.

[6] In Pari-Mutuel betting the odds change continuously based on how bets are being placed, and all bettors who take a certain bet are exposed to the final odds, regardless of the odds when they first made the bet.

efficiency of the market by regressing the actual point spread of a game against the Las Vegas line for the game. Additionally, Pankoff also tested if certain “superior analysts” exist, as the ability for certain analysts to consistently beat the spread implies that some information about the market is not accurately captured in the price of a bet. Over the 1956-1965 NFL seasons, Pankoff was unable to reject market rationality in the aggregate sense using a regression. Interestingly, Pankoff found some evidence to suggest that “superior analysts” may exist, indicating market inefficiencies which were too nuanced for his first analysis to detect.

Vergin and Scriabin (1978) were among the first to produce reasonable evidence that the NFL point spread betting market was inefficient. In their paper, Vergin and Scriabin opted to test for market efficiency by analyzing the profitability of certain betting strategies over the 1969-1974 NFL seasons. If certain betting strategies could consistently produce a profit, the market must not be efficient. Vergin and Scriabin used a number of betting strategies such as betting the underdog or trying to identify teams that recently changed in performance. The pair found some of their strategies to be both profitable and statistically significant, and concluded that there must be some inefficiency in the pricing of bets. Tryfos et al. (1984) responded to the Vergin and Scriabin (1978) paper by highlighting statistical errors which resulted in some “profitable” strategies incorrectly being labeled as statistically significant.

Zuber, Gandar, and Bowers (1985) tested the “simple” hypothesis for market rationality (that the intercept and slope of the linear regression of actual game spreads on Las Vegas spreads are jointly zero and one) using the 1983 NFL season as data. They could not reject rationality based on this regression. They also tested an ‘extreme’ alternative hypothesis that betting lines were completely unrelated to the final spread and determined that this hypothesis could not be rejected for fifteen of the sixteen weeks in the season. Zuber, Gandar, and Bowers (1985) also tested a regression model which collected data from the first half of the 1983 season and used it to predict bets with a 58.8 winning percentage over the second half of the season.

Despite the strong results found by Zuber, Gandar, and Bowers in 1985, Sauer, Brajer, Ferris, and Marr (1988) found evidence that these results did not hold when considered more generally. Firstly, Sauer et al. (1988) highlighted that by splitting the single season of data into weeks, Zuber et al. (1985) artificially increased the variance of the estimators and made it more likely that the 'extreme' hypothesis (that final spread and Las Vegas spread were completely unrelated) could not be rejected. Furthermore, Sauer et al. found that the predictive model proposed by Zuber, Gandar, and Bowers provided no additional information when compared to strictly using the Las Vegas spread as an indicator of outcome. They also found that this model did not produce profitable bets when used out-of-sample.

Later in 1988 Gandar, Zuber, O'Brien, and Russo released a paper that attempted to overcome some of the statistical shortcomings of their earlier work, in addition to providing more robust economic tests of market efficiency over the 1980-1985 seasons. Gandar et al. did not find statistical evidence to reject market efficiency. Some economic tests used bettor behavior (primarily through observation of spreads changing between open and close) and were able to consistently see profit. This paper, however, does not draw on changes in the line between open and close; data collection for historic spread movement is much harder to collect than closing lines, and thus would limit this sample.

Golec and Tamarkin (1991) attempted to address some of the area's previous statistical shortcomings by including two dummy variables for home/away and favorite/underdog in the standard test for market rationality tested by others previously (that the intercept and slope of actual spread on Las Vegas spread were 0 and 1 respectively). By including these two dummy variables Golec and Tamarkin seemed to capture bettor biases in a way that previous research had not. By applying these biases to betting strategies, Golec and Tamarkin conclude that in addition to being statistically inefficient, the NFL point spread betting market also was economically inefficient. The pair managed to highlight profitable betting strategies (bet on home underdogs) over the 1973-1987 seasons. Golec and Tamarkin did not find these same strategies profitable on college point spread betting, presumably due to a higher percentage of informed bettors as opposed to uninformed masses.

Sauer (1998) published a comprehensive literature review on wagering markets. This review covered a number of sports and gambling markets and concluded that in general these markets are efficient. Sauer's specific critiques and suggestions regarding future research in this field are noted as they occur.

Summers (2008) published an article in which he attempted to analyze the effectiveness of more complicated betting strategies, principally the effectiveness of aiming for "middles" (betting on each team at a different spread in hopes of winning both bets) on games. Summers found some evidence to suggest that this more complicated strategy could produce profits over time, but fails to reproduce any of the profits previous authors found with more simplistic betting strategies. Summers did not, however, explore simple betting strategies which used partitions in an attempt to find more statistically significant inefficiencies.

While not directly related to identifying profitable betting strategies, Levitt (2004) provided evidence via a betting tournament that suggested that book makers did not set lines in an attempt to balance their books. Levitt used data that not only included information on the spreads of games, but also on the volume of bets placed on each side of the line. Levitt claims that book makers are better at predicting outcomes than the general public, and thus are able to set spreads in such a way that the majority of bets lose, increasing book maker profits. As identified in the explanation of how books are made, this change in price must be balanced with the potential for experienced bettors to take advantage of the situation. While Levitt's results limit the ability to apply efficient market hypothesis results from the betting market to more traditional markets (where the book keepers are actually neutral), they do provide evidence of systematic bias that may be exploitable.

Paul and Weinbach (2011) expand on Levitt's (2004) evidence by analyzing the same phenomenon with a broader sample. While Levitt was constrained to a betting tournament with slightly different incentives than normal betting, Paul and Weinbach were able to obtain betting-volume data for two NFL seasons. Using regressions, Paul and Weinbach provide evidence that the volume of bets on games was systematically biased in a way that provided book makers with more profits than would be realized with an even book. In their paper Paul and Weinbach also identify some simple betting strategies that were profitable (wagering on underdogs).

As mentioned at the beginning of this literature review, the findings of past research into the efficiency of NFL point spread betting markets provides mixed signals. Many authors have found profitable betting strategies, indicating inefficiency, only to have subsequent authors provide contrasting results. One potential reason for this inconsistent nature of past research could be sample sizes that are too small to allow accurate measurement of some proposed inefficiencies. This paper aims to solve this concern by using a larger sample.

IV. Methodology

As explored by previous literature there are a number of ways to test the efficiency of the NFL betting market. In general, these tests fall under two categories: statistical tests and economic tests. Statistical tests aim to identify whether or not the predicted point spread is an unbiased, minimum variance estimator of realized point spreads, as required by EMH. Sauer (1998) and Gandar et al. (1988) represent this requirement using equation (3), where Ω represents the set of all relevant information, PS represents the Las Vegas point spread, and DP represents the actual difference in points:

$$\mathbf{E}(\mathrm{DP}\text{-}\mathrm{PS} \mid \Omega) = 0 \qquad (3)$$

The veracity of equation (3) will be analyzed in the context of the following linear prediction model:

$$\mathrm{DP} = \alpha \cdot \mathrm{I} + \beta \cdot \mathrm{PS} + \varepsilon \qquad (4)$$

Where I is a vector of ones, α and β are regression coefficients and ε is an error term. This representation of equation (3) has been used many times in the past literature to test the efficiency of the point spreads betting market; the joint hypothesis that $\alpha = 0$ and $\beta = 1$ implies market efficiency. Coefficients deviating from 0 and 1 indicate that there is some inefficiency in the market.

Alternatively, economic tests aim to identify betting strategies that produce consistent profits, as this should be impossible under the EMH. This condition is motivated by the idea that in a gambling market the book makers and betting public can not both make profits. Sauer (1998) represents this constraint using equation (5):

$$1/(2 + \tau) \leq p \leq (1 + \tau)/(2 + \tau) \qquad (5)$$

Where p represents the probability of a certain bet winning and τ represents the difference from 1 to 1 odds.[7] Using this framework one can see that in a gambling market without *vigorisih* ($\tau = 0$) the absence of a profit opportunity requires $0.5 \leq p \leq 0.5$; if p lies outside these bounds then a profit could be realized. Therefore in the case of 0 *vigorish* any value such that $p \neq 0.5$ indicates that there are profitable opportunities. When the odds are 10 to 11, as in NFL point spreads betting, $\tau = 0.1$. Under this scenario unprofitability requires that $0.476 \leq p \leq 0.524$; if p lies outside these bounds then a profit could be realized. The symmetry around $p = 0.5$ is due to the nature of bets; if the probability of a certain bet winning is 0.3, the probability of the opposite position of the same bet winning is 0.7. This bounding of unprofitable winning bet probabilities makes economic tests of market efficiency fairly straightforward. If a betting strategy can be identified to have a

[7] For example, in the case of NFL betting spreads τ would be 0.1, as book makers set odds to be 1 to 1.1 (10 to 11).

statistically significant probability of winning, *p*, which lies outside of the bounded range then the market must have some inefficiencies which can be exploited for profits.

V. Data

The data used in this paper are historical public data for the twenty one NFL seasons between years 1992-2012, and obtained from *PRO-FOOTBALL-REFERENCE.COM*. The Las Vegas lines used were compiled from PFR into a more accessible format hosted by *Aragorn Technologies*. Additionally the final scores of each team, home team designation, and dates were collected. In total this generated data for 5,192 NFL regular season observations and 231 playoff observations. For the purposes of this paper playoff games are not analyzed; the single elimination nature of playoff games, differences in player incentives (both directly and indirectly via future contracts), and modified rule set (especially with regards to home designation) seem to justify this exclusion.

As noted earlier, the spread and outcome of a game can be defined ambiguously; which frame of reference used can significantly alter the characteristics of the data and the analysis that follows. While far from an exhaustive explanation, consider the following proofs of concept: if the spread used in a game is viewed from the perspective of the favorite it has a maximum value of 0 (a “pick ‘em” game). Alternatively, when defined from the perspective of the home team the spread will take on negative values when the home team is favored and positive values when the home team is the underdog. Similarly, the final spread of the game can be defined relative to either team. One interesting caveat to these differences is that when using the favorite/underdog perspective no analysis can be done when the spread is zero (it is ambiguous which team should be labeled as the favorite). As such, any analysis from this frame of reference must exclude pick ‘em games, which can be included from the home-away perspective. Summary statistics are given in Table 1. Figure 1 and Figure 2 show the frequency distributions of the differences between actual differences in points (DP) and Vegas point spreads (PS); these graphs test equation (3). Overlaid on both of these figures is a normal distribution curve for comparison in addition to a kernel density estimate of the data. The means of DP-PS are not statistically significantly different from zero, regardless of the frame of reference used.

The difference in games per season is a result of the franchise expansions that took place during this time period. Regardless of the frame of reference used for the spreads, the mean of the Las Vegas predicted spread (PS) is extremely close to the mean of final spreads (DP). Additionally, the standard deviation of Las Vegas spreads is significantly lower than the standard deviation of final game spreads.

VI. Results of Efficiency Tests

Statistical Tests

The results of an ordinary least squares (OLS) estimation of equation (4) for each year in the data set, in addition to all years in aggregate are presented in Table 2. This estimation uses spreads from the home/away frame of reference. Before detailed analysis of these results, however, it is worth discussing the statistics presented in Table 1 again. As noted above, the mean for predicted and observed spreads were approximately even (especially when years were aggregated together), and the standard deviation of predicted spreads was lower than the standard deviation of final spreads (for every year individually and in aggregate). Both of these characteristics are required if the model suggested by equation (4) is a minimum-variance, unbiased estimator of final spreads as the EMH would require.

The coefficients of the OLS regression of DP on PS presented in Table 2 also meet the requirements set forth by efficient markets. Intercept refers to the constant, α, from equation (4), and slope refers to the coefficient on PS, β. In an efficient market the intercept, α, will be insignificantly different from zero. This condition holding implies that PS is an unbiased predictor of DP; the Las Vegas predicted spread will not be systematically above or below the actual spread. Relatedly, in an efficient market the slope, β, will be insignificantly different from one. This condition holding implies that changes in the magnitude of the actual difference in points, DP, are predicted as consistently as possible by changes in the Vegas spread, PS. If both of these conditions hold true, then the model expressed in equation (4) provides an efficient estimation which proportionally reflects changes in magnitude of DP, without being systematically biased.

When considered separately all but one of the intercept terms are insignificantly different from zero. The intercept exception is in 2006, where the intercept was shockingly large at 1.69. I offer no explanation as to why this outlier exists. The slope term from every year was insignificantly different from one. Similarly, when viewed in aggregate the intercept is insignificantly different from zero and the slope is insignificantly different from one. F-tests of the joint null hypothesis that $[\alpha,\beta]$ are jointly $[0,1]$ are also shown in Table 2. These tests confirm for all but two years, 2005 and 2006, and in aggregate that this estimation is unable to reject the null hypothesis of market efficiency represented by the joint null hypothesis. Years 2005 and 2006 reject this hypothesis at the 10% and 5% levels, respectively. While far from a comprehensive explanation of this result, it is worth noting that 2005 had the most extreme (distance from 1.0) slope coefficient and 2006 had the most extreme (distance from 0) intercept.

Despite these two years, overall the results reported in Table 2 do not contradict the hypothesis that the NFL point spread betting market is efficient. This result is not surprising; the majority of past research has indicated that the market is efficient in this aggregate sense. One weakness of this method of determining market efficiency is that equation (4) is sensitive to how the point spread is defined, and biases can exist that are not reflected in the aggregate coefficients. Sauer (1998) provides an eloquent explanation of this phenomenon (DP = final point spread, PS = Vegas predicted point spread):

> Suppose that 2/3 of the games consist of favorites playing at home, and that favorites at home are overbet on average by one point (i.e. that **E**(DP – PS) = -1 in this sub-sample). In the remaining 1/3 of the games the home team is an underdog, and suppose these teams are underbet by two points. The reciprocal nature of point differences implies that favorites playing away are overbet by two points. Thus, in the full sample, all favorites are overbet, on average, by 1.67 points. But when point differences are constructed on a home-away basis, the sample means of DP and PS are the same – the two biases posited above cancel exactly. It follows that a regression based on [equation (4)] using home-away point differences will fail to detect what is, by construction, an obvious case of mispricing. (p. 2053)

Thus, the above statistical test for efficiency is not sufficient to conclude market efficiency. Perhaps the simplest way to uncover these potentially canceling biases is to use partition-based economic tests of market efficiency.

Economic Tests

As mentioned above, a simple and effective way to identify inefficiencies that arise from complications involving variable definition in statistical models is to use partition-based economic tests. This style of test applies different betting strategies (such as bet on the favorite to cover the spread) to the games and records wins/losses. In the event a specific strategy exploits inefficiency in the market the strategy will exhibit statistically significant returns. Table 3 presents data from the NFL betting market using both the home/away and favorite/underdog perspectives, and all partitions of these orderings. The table also includes pick ‘em games, which have to be evaluated under the home/away perspective. Panel A of Table 3 tests equation (5) and finds that one betting strategy, betting on the home team in pick ‘em games, wins with a percentage outside of the bounds of unprofitability required by equation (5). As shown by the t-statistic this result is not statistically significant at any conventional level of confidence.

Despite these first partitions of games not displaying signs of any significantly profitable betting strategies, there are other ways the data can be divided which may uncover more profitable betting strategies. One hypothesis that has shown promising yet inconsistent results over previous research (Vergin and Scriabin 1978, Tryfos et al 1984) is betting on big home underdogs[8]. This strategy assumes that the betting public's inability to accurately judge home underdogs actually increases as the spread increases. When using the home/away perspective of point spreads, home underdogs can be identified by those games where the point spread is positive. Table 4 presents data on home underdogs, broken down into bins by the size of the predicted spread. The bins selected are based on the common scoring increments in football: games that are expected to be within one field goal, games that are expected to be within one touchdown and games that are expected to be decided by more than one possession.

For now, ignore the right columns that have non-zero values for "Spread Advantage". The results presented in Table 4 highlight a profound result regarding NFL point spread betting markets: large home underdogs are consistently and statistically significantly underbet. Betting on underdogs to cover the spread when they are handicapped by over 7 points results in a winning percentage of 59.69%, far beyond the 52.4% required by equation (5). This result, unlike the home team pick 'em betting strategy, is significant at the 1% level. Table 5 presents the effectiveness of this betting strategy on a year by year basis. As seen in Table 5, the number of games in which the home team is also the underdog by at least 7 points is rather rare, only 12.5 games per year on average. This low sample size per year makes it difficult to do meaningful analysis on this betting strategy without having a large number of years to pool data from. This could be a contributing factor to this market inefficiency being relatively undocumented, despite its strong statistical significance and deviation from unprofitable

[8] To clarify: betting that home teams to cover their point spread when they are large underdogs ($PS > 0$ when viewed from the home perspective).

bounding. Vergin and Scriabin (1978) were the first to highlight large home underdogs as a potential profit opportunity; however, they defined large home underdogs as those with a spread of 5 or greater, rather than 7 or greater. As such, subsequent validation of their results has used the same metric and determined this betting strategy to be unprofitable. When the same definition is applied to this data set, the win percentage drops to just 52.98%, with significance at the 39% level; this result is much less interesting.

One area also explored by Vergin and Scriabin (1978) is the effect point spread advantages might have on profitable betting strategies. Point spread advantages are situations in which one can make a bet at a spread that is markedly different from the public line. Vergin and Scriabin introduced this idea through the lens of betting syndicates, which use coordinated bettors across broad geographical regions to place bets at the most advantageous odds available to members of the syndicate. In modern times organized betting syndicates are less practical, but another force exists which allows for a more varied set of spreads to be available. With the rise of the internet, and with it online gambling, the barriers to entry for new book makers are lower, as many of the costs associated with running a gambling service are unnecessary when the service is online[9]. Due to the potential influx of online gambling books, it is not unreasonable to consider that the books may attempt to differentiate themselves with their competitors by offering mildly different prices on bets from one another. Bettors may be able to take advantage of this situation by placing their bets at certain sports books which have lower prices for

[9] There are still large barriers to entry, however. The legality of online gambling is constantly in flux as new rules and regulations are put in place in different jurisdictions. Regardless, the ability for new gambling services to enter the market has increased. Services such as Perhead white labeling (*www.perhead.com*) offer rental software designed to run sportsbooks.

certain betting positions. To consider the effects of this potential shift in the market, the Spread Advantage columns were added to Table 4.

In this hypothetical situation home underdogs who have a spread of 3 or higher are considered.[10] In these games it is assumed that, due to competition from competing gambling services, a bettor is able to find a sports book that is offering an extra 0.5, 1, 1.5, and 2 points respectively in favor of the home underdog[11]. The data shown in the additional columns of Table 4 shows the effects of this hypothetical situation. As required, win percentages increase as additional points versus the spread are gained. The statistical significant of these betting strategy winning percentages naturally increases as well.

VII. Conclusion

The NFL point spreads betting market provides an interesting analogue to more complicated financial markets. This analogous nature allows tests of economic theory such as the Efficient Market Hypothesis, which is traditionally hard to test the validity of, to be carried out with relative ease. By testing the efficiency of market pricing in the NFL point spreads betting market, conclusions can be drawn about pricing and the assumption of "informationally efficient". Early researchers in the NFL betting market used both

---

[10] When games are partitioned based on |PS| into the bins specified in Table 4, it is shown that the standard deviation of DP increases when |PS| increases. This suggests that higher spread games are harder to predict, and so book makers may be more likely to offer varied lines.

[11] Increasing the value of a spread will always benefit the team in question, regardless of whether or not they are favored. Covering the spread is calculated by the truth of this statement: DP < PS, where DP is difference in points at the conclusion of the game, and PS is the predicted point spread. Increasing PS by a spread advantage must make this condition easier to satisfy.

statistical tests, in the form of regressions, and economic tests, in the form of betting strategies, to discern the efficiency of the market. Over time, these results have been mixed, with initial results of profitable inefficiency being contradicted by later tests of the same nature.

This paper attempts to extend this analysis of simple betting strategies to modern day, when access to much larger data sets is more common. By using twenty one years of NFL game and spreads data, this paper overcomes some of the faults of earlier analyses which were limited by smaller sample sizes. The first notable finding of this paper is that the NFL point spreads betting market is efficient in aggregate. This finding confirms past research in the field still holds true when applied to more modern data sets. The second notable finding of this paper is that home team underdogs are consistently undervalued by the market, and a profitable betting strategy can be employed by betting for home underdogs to cover spreads of at least 7. This bet is successful far more frequently than is suggested by an efficient market with the structure of the NFL spreads market, and is statistically significant at the 1% level. While not an exhaustive exploration of every possible betting strategy, this result signifies that some market inefficiencies do exist, and that further exploration of the subject using more modern data still has the ability to produce meaningful new information.

This paper also briefly explores how the changing dynamics of sports books, in light of the expanding internet age, impact betting strategies. Modern software has lowered the barriers to entry in the sports book market, allowing more sports books to operate. Analysis regarding the impact of additional spread advantages (a likely byproduct of sports book competition) suggests that increased competition between sports books has the potential to increase profitable betting opportunities. One similar field of study that was not explored is the impact of lowering *vigorish*, a second likely outcome of sports book competition, on the profitability of certain betting strategies.

Further work in this field is called for, especially when considering the implementation of more complex betting strategies that leverage inefficiencies which are too subtle for the analysis in this paper to pick up on. An increased body of knowledge regarding the NFL point spread betting market will further our understanding of prices and the forces which contribute to them.

VIII. Appendix

Table 1: Summary Statistics of NFL Game Data for 1992-2012 Seasons

TABLE 1

Sumamry Statistics

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	1992-2012
Number Of Games	224	224	224	240	240	240	240	248	248	248	256	256	256	256	256	256	256	256	256	256	256	5192
Fave. Covers Spread	107	97	99	107	109	103	121	110	112	110	110	123	122	142	109	128	123	121	119	117	118	2497
Home Covers Spread	111	104	103	116	128	111	129	122	112	120	127	127	117	122	123	128	113	116	124	123	121	2407
Pick 'em Games	4	3	2	4	3	4	4	4	5	3	2	3	4	5	2	2	0	0	5	0	2	61
Pushes	3	6	5	5	5	9	10	9	7	11	5	9	5	10	5	8	6	8	5	10	4	145
Vegas Favorite Mean	-6.66	-6.00	-5.15	-5.59	-5.40	-5.13	-5.72	-5.47	-5.69	-5.20	-5.00	-5.08	-5.24	-5.15	-5.59	-5.89	-5.66	-6.70	-5.14	-5.77	-5.27	-5.61
(Standard Deviation)	(4.24)	(3.73)	(3.50)	(3.56)	(3.28)	(3.20)	(3.49)	(3.42)	(3.56)	(3.23)	(3.01)	(2.82)	(2.81)	(3.38)	(3.50)	(4.04)	(3.48)	(4.01)	(3.12)	(3.63)	(3.25)	(3.48)
Actual Favorite Mean	-6.37	-4.90	-3.80	-4.44	-4.85	-3.82	-7.05	-4.87	-5.96	-5.11	-4.20	-6.36	-5.02	-7.96	-3.67	-6.61	-6.30	-7.11	-5.33	-5.68	-5.63	-5.49
(Standard Deviation)	(14.22)	(13.64)	(12.83)	(12.78)	(13.71)	(13.97)	(12.74)	(13.68)	(13.78)	(13.30)	(13.64)	(13.76)	(13.35)	(12.93)	(13.86)	(14.24)	(14.15)	(15.00)	(14.28)	(14.44)	(14.69)	(13.84)
Vegas Home Mean	-2.86	-2.86	-2.94	-2.49	-2.90	-2.42	-2.48	-2.36	-2.57	-2.26	-2.19	-2.45	-2.52	-2.50	-2.90	-2.32	-2.69	-2.51	-2.34	-2.41	-2.19	-2.53
(Standard Deviation)	(7.39)	(6.48)	(5.50)	(6.17)	(5.63)	(5.57)	(6.25)	(6.03)	(6.23)	(5.71)	(5.42)	(5.28)	(5.42)	(5.65)	(5.93)	(6.77)	(6.08)	(7.39)	(5.57)	(6.38)	(5.80)	(6.06)
Actual Home Mean	-2.98	-2.78	-1.50	-2.03	-3.72	-2.80	-3.50	-3.06	-2.82	-2.01	-2.25	-3.55	-2.51	-3.65	-0.85	-2.87	-2.56	-2.21	-1.89	-3.27	-2.43	-2.63
(Standard Deviation)	(15.17)	(14.17)	(13.37)	(13.44)	(13.98)	(14.23)	(14.09)	(14.15)	(14.77)	(14.07)	(14.04)	(14.81)	(14.05)	(14.67)	(14.39)	(15.40)	(15.28)	(16.45)	(14.98)	(15.17)	(15.64)	(14.64)

Notes: (i) Vegas Favorite (Home) Mean refers to the mean of the Las Vegas spread, PS, from the perspective of the favorite (home) team

(ii) Actual Favorite (Home) Mean refers to the mean of the actual difference in points, DP, from the perspective of the favorite (home) team

Table 2: OLS Results of Difference in Points (DP) on Predicted Vegas Spread (PS)

TABLE 2

Regression Results for equation (4)

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	1992-2012
Intercept	-0.63	-0.37	1.28	-0.02	-1.02	-0.72	-0.70	-0.93	0.00	0.07	-0.31	-0.72	0.09	-0.51	1.691	-0.28	0.19	0.63	0.48	-0.75	0.17	-0.10
(Standard Error)	1.00	0.96	0.94	0.84	0.95	0.95	0.85	0.89	0.90	0.90	0.89	0.93	0.89	0.88	0.938	0.89	0.96	0.94	0.94	0.91	0.94	0.20
Slope	0.82	0.84	0.95	0.81	0.93	0.86	1.13	0.90	1.10	0.92	0.88	1.16	1.03	1.25	0.875	1.12	1.02	1.13	1.01	1.04	1.19	1.00
(Standard Error)	0.13	0.14	0.15	0.13	0.15	0.16	0.13	0.14	0.13	0.15	0.15	0.16	0.15	0.14	0.142	0.12	0.14	0.12	0.16	0.13	0.15	0.03
R^2	0.16	0.15	0.15	0.14	0.14	0.11	0.25	0.15	0.22	0.14	0.12	0.17	0.16	0.23	0.13	0.24	0.17	0.26	0.14	0.19	0.20	0.17
F-statistic	1.00	0.69	1.59	1.24	0.58	0.50	1.37	0.60	0.32	0.20	0.29	1.32	0.02	2.57	3.36	0.64	0.02	0.62	0.14	0.56	0.83	0.16
observations	224	224	224	240	240	240	240	248	248	248	256	256	256	256	256	256	256	256	256	256	256	5192

Notes: The F-statistic for the null hypothesis that paremeters [α, β] are jointly [0, 1]

Table 3: Score Differences and Point Spreads for NFL Games

TABLE 3

Score Differences And Point Spreads for NFL Games

A. Sample Frequencies

	Differencing Method/ Sample Partition	Games	Bets	Wins	Pushes	Wins/Bets	t-stat
A1.	Home-Away						
	All Games	5192	5047	2497	145	0.49	-4.16
	Home Favorites	3461	3358	1621	103	0.48	-4.79
	Home Underdogs	1670	1628	841	42	0.52	-0.60
	Pick 'em Games	61	61	35	0	0.57	0.78
A2.	Favorite-Underdog	5131	4986	2407	145	0.48	-7.88

B. Sample Means and Standard Deviations

	Differencing Method/ Sample Partition	DP	PS	DP-PS	t-stat
B1.	Home-Away				
	All Games	-2.63(14.64)	-2.53(6.06)	-0.10(13.33)	-0.56
	Home Favorites	-6.05(13.83)	-6.05(3.65)	0.00(13.28)	0.00
	Home Underdogs	4.36(13.78)	4.69(2.91)	-0.33(13.42)	-1.01
	Pick 'em Games	0.30(13.90)	0.00(0.00)	0.30(13.90)	0.17
B2.	Favorite-Underdog	-5.49(13.84)	-5.61(3.48)	0.11(13.33)	0.61

Notes: (i) The sample encompasses twenty one regular season NFL games from seasons 1992-2012.
(ii) Panel A: This panel lists the number of games in the specified partition. Bets refers to the number of bets in this partition which do not result in pushes. Wins refers to the number of these bets which cover the spread. Wins/Bets is the winning percentage, p, of this betting strategy. t-stat is the t-statistic for a t-test of a 1-tailed t-test that $p > 52.4\%$
(iii) Panel B: Standard deviations are listed in parentheses. The t-statistic presented tests the hypothesis that DP-PS = 0.

Table 4: Results of Betting Strategies on Home Underdogs Partitioned by PS

TABLE 4						
		Spread Advantage				
		0	0.5	1	1.5	2
0 < PS < 3						
	Number	754	-	-	-	-
	Win %	0.492	-	-	-	-
	p-value (win % < .524)	0.96	-	-	-	-
3 < PS < 7						
	Number	653	653	653	653	653
	Win %	0.48	0.49	0.51	0.53	0.55
	p-value (win % < .524)	0.99	0.96	0.76	0.47	0.11
7 < PS						
	Number	263	263	263	263	263
	Win %	0.60	0.61	0.62	0.62	0.65
	p-value (win % < .524)	0.01	0.00	0.00	0.00	0.00
Notes: (i) PS refers to the Las Vegas spread from the perspective of the home team						

Table 5: Betting on Home Underdogs of 7 or More, by Year

Table 5

Large home underdogs per year

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Number	18	13	9	17	11	11	11	12	13	10	9	7	11	10	10	19	16	27	8	14	7
Win %	0.67	0.46	0.33	0.76	0.45	0.73	0.64	0.50	0.69	0.70	0.67	0.71	0.36	0.30	0.80	0.58	0.50	0.63	0.38	0.71	0.86
P-Value (win% < 0.524)	0.11	0.66	0.86	0.02	0.52	0.09	0.24	0.52	0.12	0.14	0.21	0.17	0.84	0.91	0.03	0.32	0.57	0.14	0.78	0.08	0.03

Figure 1: The Frequency of DP-PS, Viewed from Favorite Perspective

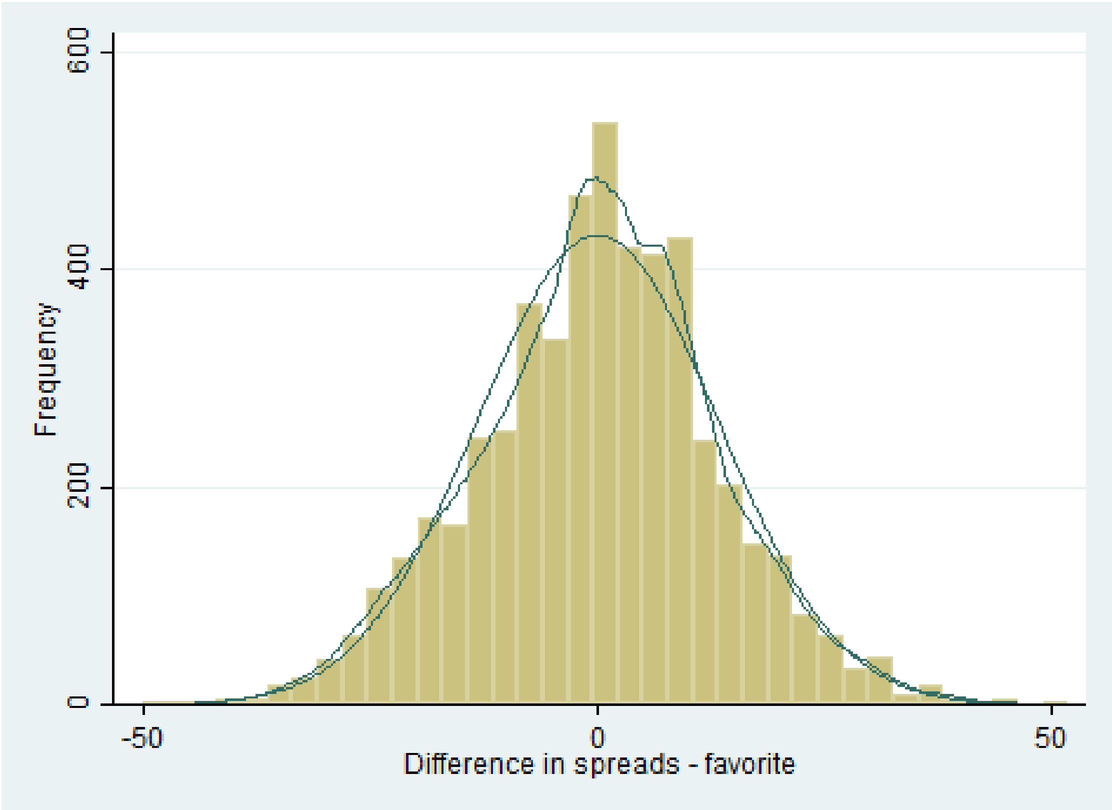
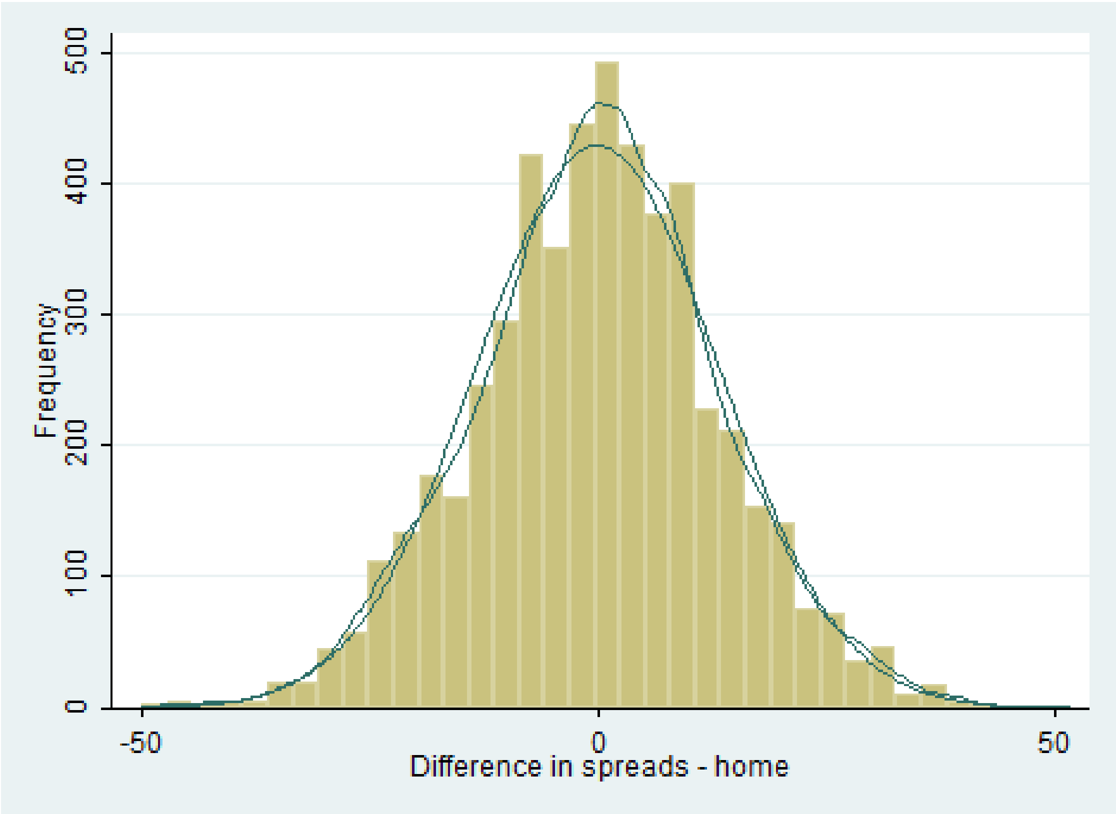


Figure 2: The Frequency of DP-PS, Viewed from Home Perspective



IX. Works Cited

Dare, William H., and S. Scott MacDonald. "A generalized model for testing the home and favorite team advantage in point spread markets." *Journal of Financial Economics* 40.2 (1996): 295-318.

Even, William E., and Nicholas R. Noble. "Testing efficiency in gambling markets." *Applied Economics* 24.1 (1992): 85-88.

Gandar, John, et al. "Testing rationality in the point spread betting market." *The Journal of Finance* 43.4 (1988): 995-1008.

Gandar, John M., Richard A. Zuber, and Reinhold P. Lamb. "The home field advantage revisited: a search for the bias in other sports betting markets."*Journal of Economics and Business* 53.4 (2001): 439-453.

Golec, Joseph, and Maurry Tamarkin. "The degree of inefficiency in the football betting market: Statistical tests." *Journal of Financial Economics* 30.2 (1991): 311-323.

Gray, Philip K., and Stephen F. Gray. "Testing market efficiency: Evidence from the NFL sports betting market." *The Journal of Finance* 52.4 (1997): 1725-1737.

Levitt, Steven D. "Why are gambling markets organised so differently from financial markets?*." *The Economic Journal* 114.495 (2004): 223-246.

Pankoff, Lyn D. "Market efficiency and football betting." *Journal of Business*(1968): 203-214.

Paul, Rodney J., and Andrew P. Weinbach. "NFL bettor biases and price setting: further tests of the Levitt hypothesis of sportsbook behaviour." *Applied Economics Letters* 18.2 (2011): 193-197.

Sauer, Raymond D. "The economics of wagering markets." *Journal of economic Literature* (1998): 2021-2064.

Sauer, Raymond D., et al. "Hold your bets: Another look at the efficiency of the gambling market for National Football League games." *The Journal of Political Economy* (1988): 206-213.

Summers, Michael R. "Beating the book: are there patterns in NFL betting lines?." *UNLV Gaming Research & Review Journal* 12.1/2 (2008): 43-52.

Tryfos, Peter, et al. "The profitability of wagering on NFL games." *Management Science* 30.1 (1984): 123-132.

Vergin, Roger C. "Overreaction in the NFL point spread market." *Applied Financial Economics* 11.5 (2001): 497-509.

Vergin, Roger C., and Michael Scriabin. "Winning strategies for wagering on National Football League games." *Management Science* 24.8 (1978): 809-818.

Woodland, Linda M., and Bill M. Woodland. "Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market." *The Journal of Finance*49.1 (1994): 269-279.

Woodland, Linda M., and Bill M. Woodland. "Market efficiency and profitable wagering in the National Hockey League: can bettors score on longshots?."*Southern Economic Journal* (2001): 983-995.

Zuber, Richard A., John M. Gandar, and Benny D. Bowers. "Beating the spread: Testing the efficiency of the gambling market for National Football League games." *The Journal of Political Economy* (1985): 800-806.